Computational Insights into Reactions of Controlled Molecules

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Uxía Constanza Rivero González

aus Spanien

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von
Prof. Dr. Stefan Willitsch und Prof. Dr. Markus Meuwly.

Basel, den 15. Oktober 2019

Prof. Dr. Martin Spiess

Dekan

Abstract

The regio- and stereoselective Diels-Alder reaction in which a diene reacts with a dienophile to form a cyclic product is widely used in synthetic organic chemistry. Many studies have aimed to unveil this mechanism and its dependence on the geometrical and electronic characteristics of the reactants.

Recent advances in molecular-beam experiments allow now conformational separation of isomers by electrostatic deflection of a molecular beam based on their different dipole moments. Hence, the separation of the conformers of a diene as well as the separation of different rotational states of small molecules is possible as long as the experimental requirements are met.

Here, the Diels-Alder reaction between conformationally selected 2,3-dibromo-1,3-butadiene (DBB) and supersonically cooled maleic anhydride (MA) is computationally studied both in its neutral and cationic variant as the first step of a combined experimental and theoretical study. Density functional theory (DFT) calculations show that the neutral reaction is concerted while the cationic reaction can be concerted or stepwise. RRKM calculations suggest that, under typical single-collision conditions, the neutral Diels-Alder product may reform the reactants and the cationic product will most likely eliminate CO_2. Reactive atomistic simulations on the neutral reaction indicate that rotational energy is crucial to drive the system towards the transition state in addition to collision energy. Comparison with the reaction of butadiene and MA shows that the presence of bromine substituents in the diene accentuates the importance of rotational excitation to promote the reaction. At the high total energies at which reactive events are recorded, the reaction is found to be direct and mostly synchronous. Reactive dynamics on the cationic reaction show that rotational energy promotes the reaction and that the recorded events are direct and mostly asynchronous.

The Diels-Alder reaction between DBB and propene has also been studied as an alternative to the reaction with MA. The neutral reaction is predicted to be concerted but with higher activation energies than the reaction of DBB with MA. The cationic variant can be either barrierless, concerted or stepwise with submerged barriers.

The proton-transfer reaction between N_2H^+ and H_2O has been studied and found to be barrierless which supports experimental findings.

Acknowledgements

Y quizá hayas andado el camino ya, cuando mires atrás
Si estás atrapado en las sombras, aguarda, aguarda.
Del lodo crecen las flores más altas, más altas.

Lodo - Xoel López

First of all I would like to thank Stefan Willitsch and Markus Meuwly for believing in me four years ago after a short skype conversation and giving me the opportunity and guidance to work in this project and come to Basel. If things would have turned out differently I would not be who I am today.

Thanks a lot to Jesús Flores and Sonia Álvarez Barcia for introducing me to the world of computational chemistry and not forgetting me once I left.

In the last four years I had the opportunity to work with many people and to be at the intersection of two research groups which has enriched me doubly. Thanks to all current and past members of the Willitsch and Meuwly groups for all the times we spent together. Also many thanks to Philip, Grisha ang Georg for giving me the opportunity to practice my German and being really nice to me. Special thanks to Anatoli for taking care of us with coffee and cookies. Thanks to Pascal and Alex for their help with German (very much appreciated).

For the scientific part, I would like to thank Ian Rouse, Sebastian Brickel and Oliver Unke for all the help and questions that they have answered. Thanks to Patrik Stranak, Sebastian Brickel (again), Thomas Kierspel and Ardita Kilaj for proofreading parts of this manuscript.

Ardita deserves a special mention. She is the first woman with whom I have practiced sorority and we need more women in science.

I would also like to thank my Spanish family: Nacho, Aroa, Olai, Hele, Pablo, Jorge, abuela, abuelo, Pablo, Sonia, Fer, Maria, Eva, Miguel, Bruno, Nico, Mario, Lucas, Sonia, Santiago because, even from 2000 km away, you always pick up when I call, you make sure that I know that you are proud of me and that I can count on you and you are very happy whenever I go visit.

Thanks also to my Swiss family: Maridel, Eva, Nora, Pascal and Kevin for your close support and nice times and for being guilty of me wanting to stay in Basel.

Finally, thank you Michael for your support, your patience, your humor, your empathy and for all the time that you choose to spend with me. Thanks to you I survived the last month of writing and Basel felt like home very early, every time that I was on your red sofa.

This thesis is dedicated to Ramón Rivero González because I hope I can be more and more like him and because he was right when he told me that Switzerland was going to treat me “como a una reina” (even if I am republican) and because I am very lucky and happy to be his granddaughter.

Contents

Chapter 1

Introduction

In 2013, Martin Karplus won the Nobel prize (together with Michel Levitt and Arieh Warshel) "for the development of multiscale models for complex chemical systems" [1]. Karplus *et al.* were the first ones to perform reaction dynamics. They developed, in the 1960s, a quasiclassical procedure for one of the simplest reactions that we can think of: $H + H_2$ [2–4]. The field has since then evolved and we can now simulate more complex reactions, proteins or even cellular processes [5, 6]. In the real world, particles have kinetic energy and their motion will affect their physical and chemical behavior. Reaction molecular dynamics allows us to "see" (simulate) molecules and atoms in motion. For this reason, it is a valuable tool to predict experimental results.

1.1 Crossed Molecular Beam and Ion Trap Experiments

When performing calculations on chemical systems, there is full control over the degrees of freedom of the system that can be initialized in the desired geometry and with the most convenient energy distribution to study the process of interest. Experimentally, this is more challenging and many efforts are being made towards gaining more control for the preparation of the systems of study.

In order to understand the fundamental aspects of chemical systems, we need to start with the simplest ones and build up from there. Gas-phase systems in which there is minimal interaction between the individual molecules are thus a

good starting point. If we have a reservoir filled with a molecular gas at a temperature T, the individual molecules will have a certain distribution of translational, vibrational and rotational energy. Limited control over the total kinetic energy of the system can be achieved by heating up or cooling down the whole reservoir but the distribution of energy will still be thermal and the molecules will still move in random directions.

If we want to gain more control over the translational energy of the particles in the reservoir such that, for example, we can direct the molecules towards a target in order to study scattering processes, a molecular beam can be generated. The first beam composed of neutral atoms was produced in 1911 by Dunoyer [7] shortly after the invention of high speed vacuum pumps [8–10]. A molecular beam is generated when gas contained in a high-pressure reservoir expands through a small orifice (called nozzle) into a low-pressure region. In a collimated beam, that can be generated with specific geometries of nozzles and skimmers, all particles travel in the same direction with a small perpendicular spread.

Today, molecular beams are often generated by supersonic expansion: a gas in a reservoir is expanded into vacuum through a nozzle much larger than the mean free path of the particles (the average distance that particles travel between two collisions) such that many collisions are produced as the gas leaves the reservoir. These beams are usually composed of a carrier gas seeded with the molecule or atom of study. In this way, most of the particles are carrier-gas atoms (with no rotational or vibrational degrees of freedom) so the seeded molecules will mostly inelastically collide with those atoms and their vibrational and rotational degrees of freedom will be cooled down. The rotational degrees of freedom will be more efficiently cooled down because there is quite a mismatch between the energy transfered in the collisions and the vibrational energy levels that are farther separated than the rotational levels. An advantage of using seeded molecular beams is that the molecules or atoms of study are, in practice, isolated from each other.

For the study of a reaction between two species, two beams can be designed to interact with each other. There are different ways to do it, two examples are merged beams experiments [11] and crossed molecular beam experiments [12] like the one depicted in figure 1.1. The collision energy in a crossed molecular beam experiment can be varied by modifying the angle of the colliding beams [13], the

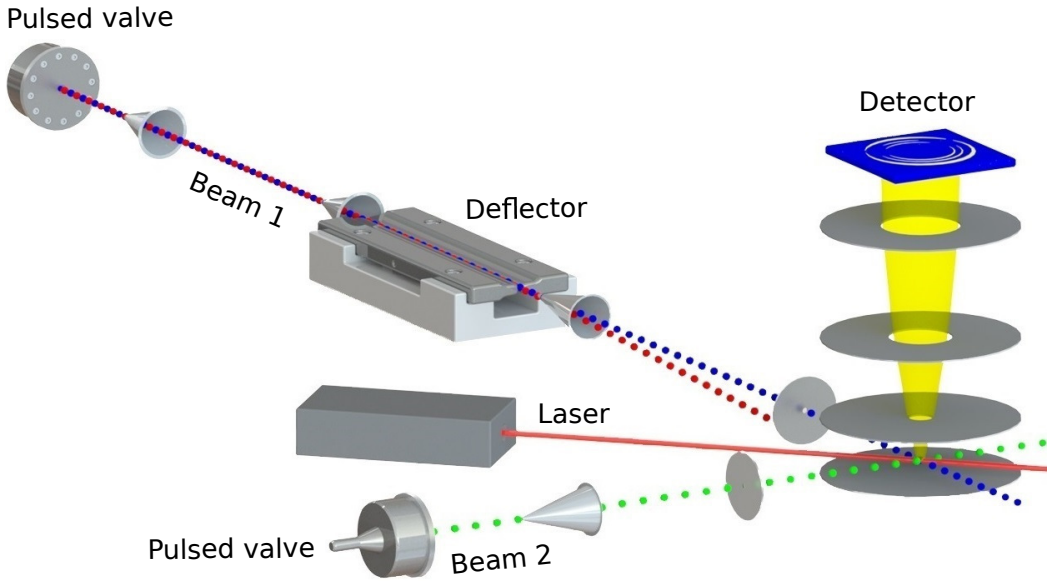


Figure 1.1: Schematic of a crossed molecular beam setup. Two pulsed molecular beams are generated with the pulsed valves. The first molecular beam (beam 1) is composed of two conformers of the same molecule (represented as blue and red balls) and passes through a deflector where an inhomogeneous electric field spatially separates these two conformers. The blue conformer collides with the second beam (beam 2) represented by green balls. A laser ionizes the products that are extracted and detected by a velocity map imaging spectrometer [19]. Figure credit: Ludger Plönes.

temperature of the valves [14] or by exchanging the carrier gas since the velocity of the beam depends on the mass of the particles that compose it. Besides, the longitudinal velocity of a beam can be reduced by using decelerators [15–18].

A conformer or state selected molecular beam can be obtained by modifying the transverse velocity of the beam with inhomogeneous electric and magnetic fields [20–23]. Polar molecules in an electric field experience Stark shifts that depend on the strength of the electric field and on the effective dipole moment of the molecule. Since different rotational levels have different effective dipole moments, a spatial separation of rotationally selected molecules can be achieved [24, 25]. This technique can also be used for the separation of conformational isomers that differ in their dipole moments (see the deflector in figures 1.1 and 1.2) [26]. In this way, only the selected part of the beam will be used for the experiment.

Traps can be used to confine particles [27–29]. Different kinds of traps exist: neutral atoms can be trapped with magneto optical traps [27] where an inhomogeneous magnetic field in combination with a spatial superposition of counter propagating laser beams (optical molasses) is used to cool down the particles

through position-dependent Doppler cooling. In this technique, a laser is tuned such that an electronic transition of the atom of interest can be excited. The atom will later emit this photon in a random direction to return to the initial state. By red detuning the laser in such a way that only atoms traveling towards the direction of propagation of the laser absorb photons, a net force will be created resulting in the deceleration of those atoms (the absorbed photons will always come from the same direction while the emitted ones will leave in random directions). The inhomogeneous magnetic field gradually splits the energy levels of the trapped particle as it moves out of the center of the trap such that the cooling force depends on the position of the particles in the trap.

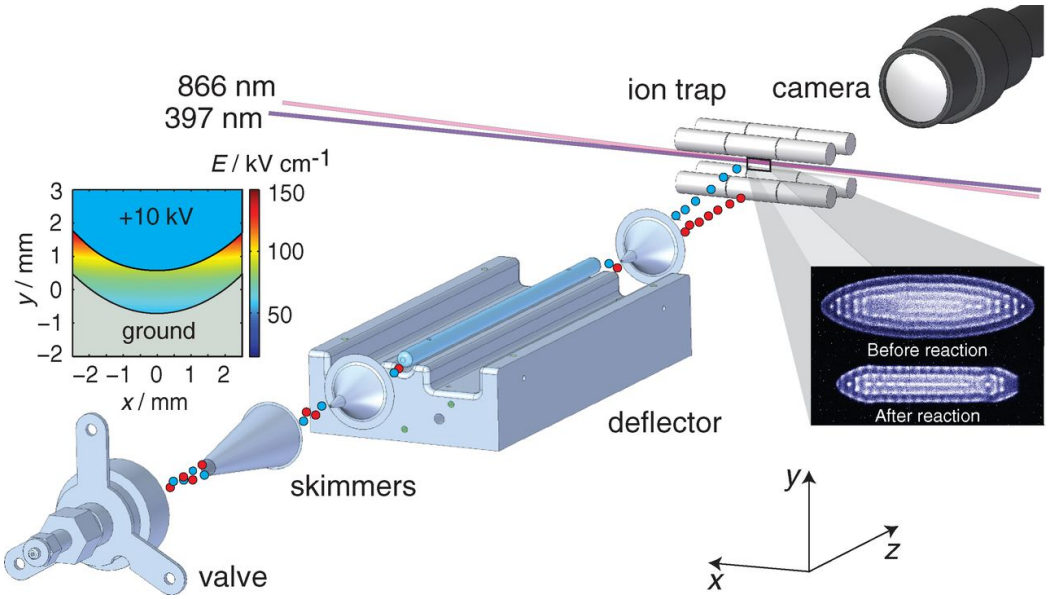


Figure 1.2: A pulsed molecular beam of two conformers of the same molecule (represented as blue and red balls) is generated and passes through a deflector that separates them spatially with an inhomogeneous electric field (see inset in the left). After the deflector, the selected part of the beam is directed into an ion trap and collides with sympathetically cooled ions that sit inside a Ca^+ Coulomb crystal (in the inset in the right). The lasers with wavelengths 866 and 397 nm are used to Doppler cool the Ca^+. The reaction is examined through the changes in the Coulomb crystal recorded by the camera [26].

Charged particles can be trapped in quadrupole radio-frequency ion traps in which a potential well is generated with a dynamic electric field. A static electric field would only create a minimum in two (confining) directions and a maximum in the third (anti-confining) direction (Earnshaw's theorem). In order to generate a trapping potential, the ions are confined with a static DC field in one of the directions and with an electrodynamic field in the other two directions. The dynamic field switches fast between the confining and anti-confining direction

such that the ions cannot escape the trap [28]. The trapped ions can be laser cooled as explained above. Ideally, the excited state will relax back directly to the ground state such that the cooling cycle can be repeated many times. However, even for atoms, it is sometimes experimentally challenging to have what is called a closed cooling cycle since there is a certain probability that the relaxation of the atom leads to a state different to the ground state. When this happens, additional lasers are needed to pump the population from those different states back to the excited state such that the ions can continue to be cooled. Ca^+ (and other alkaline-earth ions) is commonly used since two lasers are enough to have a closed cooling cycle. Simple molecules can also be laser cooled but require much more complex cooling schemes [30].

Very cold, trapped ions can form ordered structures called Coulomb crystals [31]. Cationic molecules or atoms that cannot be Doppler cooled can be trapped inside a Coulomb crystal and be sympathetically cooled by the ions conforming the Coulomb crystal through elastic collisions (see inset in figure 1.2) which makes this technique much more flexible. Reactions can be studied by directing a molecular beam towards a Coulomb crystal and followed by the changes happening to the crystal [25, 26].

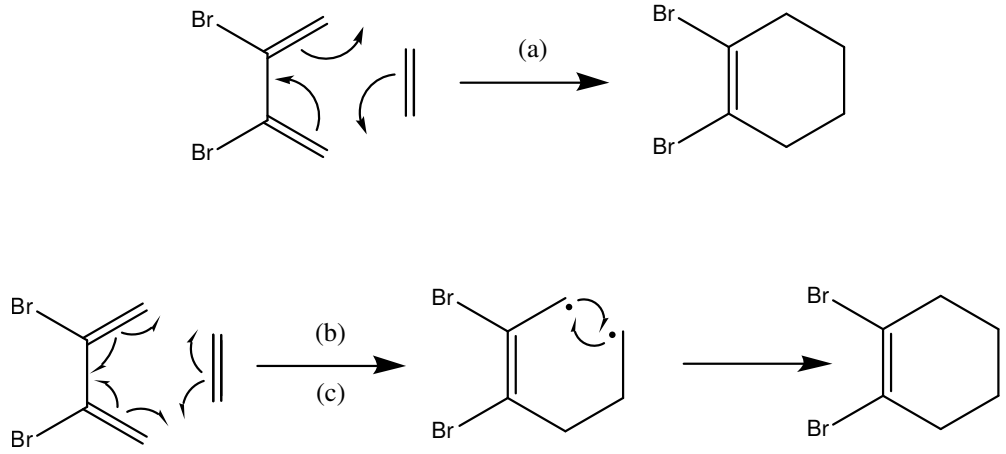
Hybrid traps, that combine a magneto optical trap with a linear quadrupole trap, have also been developed and can be used to study reactions between ions and neutral particles since they can trap, in the same region of space, both types of particles [32–35].

The influence of vibrational energy can be studied by exciting specific vibrations of molecules [36, 37]. Rotational excitation can also be achieved [38–43].

In this thesis, calculations are performed to aid the experimental study of conformational selective reactions. One of these reactions is the Diels-Alder reaction [44] where only the s-*cis* conformer is believed to react when the reaction is concerted. If the conformers of the diene have different dipole moments, a beam of conformationally selected s-*cis* or s-*trans* diene molecules can be obtained with the techniques discussed above (see figures 1.1 and 1.2). The selected conformer will then collide with the dienophile and prove whether the s-*cis* conformer is really the only one than can yield the Diels-Alder product.

1.2 Diels-Alder Reactions *

The regio- and stereoselective Diels-Alder reaction in which a diene reacts with a dienophile to form a cyclic product is widely used in synthetic organic chemistry [44, 45]. In this reaction, two σ bonds and one π bond are formed from three π bonds as depicted in scheme 1.1.



Scheme 1.1. Diels-Alder reaction between 1,3-butadiene and ethylene: (a) Concerted mechanism, (b) and (c) stepwise mechanism with a short-lived and a long-lived intermediate, respectively.

Many experimental and theoretical studies have aimed at unveiling the Diels-Alder mechanism at a molecular level and its dependence on the geometric and electronic characteristics of the reactants [46–57]. Since two bonds are formed, questions about concertedness and synchronicity render this reaction also interesting from a theoretical point of view.

The concertedness of a mechanism is determined by the topology of the potential energy surface (PES) [58]. A concerted mechanism occurs when the PES exhibits only one transition state (TS) between reactants and products so that the process takes place in a single step. A mechanism will be stepwise (taking place in two or more elementary steps) when the system has to overcome at least two TSs separated by an intermediate species (a local minimum on the PES) to evolve from reactants to products.

*This section is based on the introductions of the papers *A computational study of the Diels-Alder reactions between 2,3-dibromo-1,3-butadiene and maleic anhydride.* U. Rivero, M. Meuwly, S. Willitsch, Chemical Physics Letters 683 (2017) 598 and *Reactive Atomistic Simulations of Diels-Alder reactions: the Importance of Molecular Rotations.* U. Rivero, O. T. Unke, M. Meuwly, S. Willitsch, Journal of Chemical Physics 151 (2019) 104301.

The time elapsed between the formation of the first and the second bond defines the synchronicity of the process [58]. It is usually thought that symmetric TSs give rise to synchronous processes in which both bonds are formed at the same time while asymmetric TSs lead to asynchronous processes in which first one bond is formed and then the second one follows. This definition of synchronicity has recently been challenged by some authors who argue that it should not be defined from a geometric but from a dynamic point of view since the connection between spatial quantities and temporal concepts may not always hold [48, 59]. A synchronous process is concerted while an asynchronous one can be concerted or stepwise depending on the absence of a (stable) intermediate.

There has been a long-standing discussion about synchronicity and concertedness of Diels-Alder reactions which is not yet resolved [46, 49]. The picture that organic chemistry textbooks usually give is that it is a concerted, synchronous reaction governed by the Woodward-Hoffmann rules that involves an aromatic TS [45, 60]. However, experiments and calculations show that this is not that simple in many cases. In principle, one can think of three possible mechanisms (see scheme 1.1): (a) synchronous concerted, (b) stepwise with a short lived intermediate whose lifetime is not long enough for the system to rotate around a C-C bond, and (c) stepwise with a long lived intermediate. Note that when the system cannot rotate around a C-C bond, as is the case in (a) and (b), the reaction is stereo-selective and only the s-*cis* conformer of the diene will yield the cyclic Diels-Alder product. On the contrary, mechanism (c) is not stereo-selective and the s-*trans* conformer of the diene could also in principle yield a Diels-Alder product.

Zewail and his group have performed different experiments involving retro Diels-Alder reactions in which they detected intermediates [50, 51]. This suggests that the reaction must happen, at least partially, in a non-concerted fashion in these cases. However, since the experiment is started with the excitation of the Diels-Alder product, electronic excited states play a role in the dynamics [61]. There also exist several studies involving kinetic isotope effects that compare experimental and calculated results in order to determine if the reaction is concerted or stepwise. The picture emerged that neutral Diels-Alder reactions tend to take place in a concerted way while ionic Diels-Alder reactions usually happen in a stepwise manner [52–54].

Computational studies of Diels-Alder reactions are challenging since the relevant systems are comparatively large and especially sensitive to the choice of method and basis set [62, 63]. The reaction between butadiene and ethene (the simplest Diels-Alder reaction), see scheme 1.1, has been studied at various levels of theory [46, 55, 64–66]. For this symmetric reaction, *ab initio* and density functional theory (DFT) calculations favor a concerted, synchronous mechanism (as suggested by the symmetry of the system), but multireference methods are needed to accurately describe the activation energy and the enthalpy of reaction as compared to the experimental values [64, 65, 67]. DFT methods can also provide good results, but due to the wide range of resulting energies the choice of functional is not trivial [65, 66].

The complexity is higher when studying asymmetric reactions due to the fact that the choice of the method influences the degrees of synchronicity and concertedness of the processes and the (a)symmetry of the TS [56]. Moreover, there seems to be a correlation between the asymmetry of the TS structure and the rate constant of the process, where asymmetry refers to the difference in the lengths of the newly formed σ bonds at the TS. The more asymmetric the TS geometry, the faster the reaction [47, 56]. Diels-Alder reactions are usually activated by electron-withdrawing groups in the dienophile and electron-rich groups in the diene although the opposite situation, in which the electron-rich groups are in the dienophile and the electron-withdrowing ones in the diene, is also possible [45].

In general and compared to experimental values and high level theoretical calculations, Møller-Plesset second order perturbation theory (MP2) was found to underestimate the activation energies because of an overestimation of electron correlation [65]. Complete active space self consistent field (CASSCF) calculations give values close to Hartree-Fock (HF) due to its lack of dynamic correlation, which causes an overestimation of the height of the reaction barrier [64, 65]. It is thus necessary to use complete active space second order perturbation theory (CASPT2) to obtain accurate energies [65]. On the other hand, DFT methods predict a wide range of activation energies and reaction enthalpies and the widely used B3LYP/6-31G* approach tends to overestimate the activation energies while underestimating the reaction enthalpies [63, 66, 68]. Moreover, it seems that increasing the basis set does not improve the quality of the results and it can even make them worse. A medium-size basis set, such as 6-31G*, has thus been rec-

ommended for the study of Diels-Alder reactions when used in combination with DFT methods [62, 68, 69].

Cationic Diels-Alder reactions (polar cycloadditions) are often faster but still show a high degree of stereoselectivity [70–72]. There have been some studies on the conservation of orbital symmetry to try to construct rules analogous to the Woodward-Hoffmann rules widely used for the neutral reactions [72–74]. Wiest and Donoghe proposed a model in which the electronic state symmetry must be conserved throughout the reaction [57]. Gas-phase experiments on the cationic Diels-Alder reaction between butadiene and ethene have been unable to isolate the Diels-Alder product. As no efficient deactivation was possible in the gas phase, the product must have fragmented under the experimental conditions [75]. Subsequent computational studies explored the possible fragmentation pathways of the Diels-Alder product in order to interpret the experimental findings [76, 77].

In summary, these results corroborate the picture that neutral reactions usually occur in a concerted fashion while in cationic systems a non-concerted mechanism is favored [46, 57]. However, the border between asynchronous, concerted and stepwise mechanisms is not yet clear [52, 53, 56]. It has also been argued that both concerted and stepwise mechanisms can be present for the temperatures at which these reactions are usually performed (around 500 K and above) due to the energetic proximity of both pathways in many systems [48, 50, 51] making dynamical studies crucial for the exploration of these reactions. Besides, when a stepwise process takes place, the competition between the closure of the ring and the isomerization of the intermediate state needs to be studied in order to determine whether the reaction will be stereo-selective. To the best of our knowledge, all theoretical studies on the dynamics of Diels-Alder reactions have started from TS-like structures [48, 59, 78–80] or have used steered dynamics to drive the reaction [81]. These procedures will most likely bias the final result and do not allow the direct calculation of reaction rates. Gas-phase reactive molecular dynamics simulations starting from an equilibrated ensemble of a statistically significant number of initial conditions, on the other hand, have recently been shown to provide molecular-level details into reactions relevant to atmospheric chemistry [5, 82] and reactions in the hypersonic regime [83, 84].

1.3 Studied Systems

Even though most reactions take place in the solution phase, from an experimental perspective, the most precise data on reaction mechanisms and dynamics can be gained from gas-phase studies performed under single-collision conditions. The mechanism in solution phase may differ since the environment can play a big role but this is not the scope of this work.

As the progress in molecular-beam experiments allows the probing of ever-larger systems under precisely defined conditions [26, 85], the open questions pertaining to the mechanistic details of Diels-Alder reactions become an attractive target of study. The first studied system (in chapters 3, 4 and 5) is the Diels-Alder reaction of 2,3-dibromo-1,3-butadiene (DBB) as the diene and maleic anhydride as a dienophile both in its neutral and cationic variants. We use this system as a prototype to probe general mechanistic aspects of Diels-Alder reactions in the gas phase. DBB has been chosen because it is a generic, activated diene which fulfills the experimental requirements for conformational separation of its isomers by electrostatic deflection of a molecular beam [26] thus enabling the characterization of conformational aspects and specificities of the reaction. MA is a widely used, activated dienophile which due to its symmetry simplifies the possible outcomes of the reaction.

The second system studied (in chapter 6) is the Diels-Alder reaction of DBB with propene as an alternative to the reaction of DBB with MA again both in its neutral and cationic variants.

Finally, in chapter 7, the reaction between nuclear-spin-selected water and sympathetically cooled N_2H^+ is computationally studied to aid with the interpretation of experimental results.

Chapter 2

Theoretical Background

The different methods used for the completion of this thesis are summarized here. The chapter starts with the notion of potential energy surfaces and how to study them, continues with kinetic studies and finishes with classical molecular dynamics.

2.1 The Study of Potential Energy Surfaces

A potential energy surface (PES) is a multidimensional function that describes the potential energy of a system in terms of the nuclear coordinates. It arises from the Born-Oppenheimer approximation that separates the electronic motion from the nuclear motion on the basis that electrons adjust almost immediately to the new position of the nuclei when the latter move. In this way, the nuclear coordinates can be treated as parameters for the calculation of the electronic energies. At a certain configuration of the nuclei $\mathbf{R}$ the energy of the system in atomic units is given by:

$$E(\mathbf{R}) = V(\mathbf{R}) + E_{e,k}(\mathbf{R}) = \sum_{A} \sum_{B>A} \frac{Z_A Z_B}{R_{AB}} + E_{e,k}(\mathbf{R}), \tag{2.1}$$

where $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$ and Z_A is the atomic number of atom A. The electronic energy for the electronic state k ($E_{e,k}$) is then obtained by solving the non-relativistic, time-independent Schrödinger equation for a given nuclear configuration:

$$\hat{H}(\mathbf{r};\mathbf{R})\ \Psi_k(\mathbf{r};\mathbf{R}) = E_{e,k}(\mathbf{R})\ \Psi_k(\mathbf{r};\mathbf{R}), \tag{2.2}$$

where Ψ_k is the wave function for the electronic state k and $\hat{H}$ is the Hamiltonian of an N-electrons system that contains the electronic kinetic energy, nuclear-electron attraction and electron-electron repulsion:

$$\hat{H}(\mathbf{r};\mathbf{R}) = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_i \sum_A \frac{Z_A}{r_{iA}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}}, \tag{2.3}$$

where i and j run over electrons and A runs over nuclei. The electronic energy $E_{e,k}$ added to the nuclear repulsion term (equation (2.1)) defines the PES for the electronic state k.

2.1.1 Electronic Energy Calculations

The Schrödinger equation is far too complex for most systems to be solved. Thus certain approximations need to be made. The most important groups of methods to solve the Schrödinger equation are *ab initio* methods, semiempirical methods, density functional theory methods and force fields. Recently, machine learning algorithms are starting to be applied to this problem.

2.1.1.1 Ab Initio Methods

One of the first methods proposed to solve the Schrödinger equation for complex systems was the Hartree-Fock method [86, 87]. Since it introduces many concepts necessary to understand the methods that came after it, it will be explained in detail here.

For an ideal system where electrons do not interact with each other, the electron-electron repulsion term in equation (2.3) can be dismissed. The Hamiltonian is thus separable and the total electronic wave function can be written as the product of one-electron wave functions (equation (2.4)) which is known as the Hartree product:

$$\Psi_{HP}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\cdots\chi_N(\mathbf{x}_N), \tag{2.4}$$

where $\chi_i(\mathbf{x}_i)$ is the spin orbital of electron i which is a spatial orbital multiplied by a spin function.

Equation (2.4) is not a good representation for electrons, which are fermions, since the wave function should be antisymmetric with respect to the exchange of two electrons. In order to overcome this problem, Slater introduced determinants, named after him, for constructing appropriate wave functions:

$$\Psi_{SD} = |\chi_1\chi_2\cdots\chi_N\rangle = \frac{1}{\sqrt{N}}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}. \tag{2.5}$$

The use of this representation for the wave function implies that, by construction, the electrons are indistinguishable and move independently from each other (lack of dynamic correlation) in a mean field generated by the rest of the electrons. However, because of antisymmetrization, the motion of two electrons with parallel spins is correlated (exchange correlation) while the motion of electrons with opposite spins is uncorrelated. The Slater determinant in equation (2.5) has the same number of electrons and spin orbitals and thus all spin orbitals will be occupied.

According to the variational principle, the best orthonormal spin orbitals $\{\chi_i\}$ for a given Slater determinant are those that minimize the energy. For a closed shell system:

$$E_{HF} = \langle\Psi_{SD}|\hat{H}|\Psi_{SD}\rangle. \tag{2.6}$$

Taking a closer look to the previous equation:

$$E_{HF} = \sum_i \langle\chi_i|\hat{h}_i|\chi_i\rangle + \frac{1}{2}\sum_i\sum_j(\langle\chi_j|\hat{J}_i|\chi_j\rangle - \langle\chi_j|\hat{K}_i|\chi_j\rangle), \tag{2.7}$$

where $\hat{h}_i$ is the one-electron operator that accounts for kinetic energy and nuclear-electron attraction,

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_{1A}}, \tag{2.8}$$

and $\hat{J}_i$ and $\hat{K}_i$ are the so called Coulomb and exchange operators that describe electron-electron interactions:

$$\hat{J}_i|\chi_j(\mathbf{x}_2)\rangle = \langle\chi_i(\mathbf{x}_1)|\frac{1}{r_{12}}|\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\rangle, \tag{2.9}$$

$$\hat{K}_i|\chi_j(\mathbf{x}_2)\rangle = \langle\chi_i(\mathbf{x}_1)|\frac{1}{r_{12}}|\chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)\rangle. \tag{2.10}$$

The equation that determines the best spin orbitals for a single Slater determinant wave function using the variational principle is the Hartree-Fock equation

$$\hat{F}_i\chi_i = \epsilon_i\chi_i, \tag{2.11}$$

where ϵ_i is the energy of the i-th spin orbital χ_i and $\hat{F}_i$ is the Fock operator of the form

$$\hat{F}_i = \hat{h}_i + \sum_j(\hat{J}_j - \hat{K}_j). \tag{2.12}$$

Due to the fact that the solution for each spin orbital depends on the solutions of all the others, the Hartree-Fock equation must be solved iteratively in a self consistent field (SCF) manner. However, equation (2.11) is still too complicated to solve in many cases and only numerical methods can be employed [87]. Luckily in 1951, Roothaan showed that the Hartree-Fock equation can be converted (for a closed shell system) into a set of equations that can be solved by standard matrix techniques if the spatial part of the unknown orbitals (also known as molecular orbitals) were expanded through a set of K known basis functions (also referred to as atomic orbitals):

$$\chi_i = \sum_{\mu}^{K} C_{\mu i}\phi_\mu, \tag{2.13}$$

where the coefficients $C_{\mu i}$ are now the only unknowns. Equation (2.11) may then be rewritten in a matrix form as:

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon}, \tag{2.14}$$

where

$$F_{\mu\nu} = \langle\phi_\mu|\hat{h}_i|\phi_\nu\rangle + \sum_{\lambda}^{K}\sum_{\sigma}^{K} P_{\lambda\sigma}(2\langle\phi_\mu\phi_\sigma|\frac{1}{r_{12}}|\phi_\nu\phi_\lambda\rangle - \langle\phi_\mu\phi_\sigma|\frac{1}{r_{12}}|\phi_\lambda\phi_\nu\rangle), \tag{2.15}$$

$$S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle, \tag{2.16}$$

$$P_{\lambda\sigma} = 2\sum_{i=1}^{occ.} C^*_{\lambda i}C_{\sigma i}. \tag{2.17}$$

S is the overlap matrix and **P** the density matrix that considers only the occupied orbitals (occ.) since they are the ones that contribute to the Hartree-Fock energy. If the set of basis functions ϕ_μ is complete, the solution given by equation (2.14) is identical to that of equation (2.11) and any kind of basis functions could be used. In practice this cannot be done due to the high computational cost that it would imply (see section Basis Sets for more information). If there are more basis functions than electrons, there will be more orbitals than needed for the allocation of all electrons and thus some orbitals will be empty. These are called virtual orbitals.

It is important to keep in mind that equation (2.14) will not give the exact solution of the Schrödinger equation but the best possible solution for a single Slater determinant representation of the wave function. As discussed above, using a single Slater determinant for the representation of the wave function (equation (2.5)), comes with some dynamic correlation problems. Moreover, this description is not flexible enough for some systems that have nearly degenerate states (static correlation). In order to create a better wave function, a linear combination of Slater determinants can be used [87]:

$$\Psi = c_0\Psi_{HF} + c_1\Psi_1 + c_2\Psi_2 + \ldots, \tag{2.18}$$

where the coefficients c_i are the weights of the different Slater determinants and ensure normalization. The new Slater determinants differ from Ψ_{HF} by n spin orbitals. n defines the degree of excitation of the new Slater determinant. As said before, if the number of basis functions in equation (2.13) is higher than the number of orbitals needed to allocate the electrons of the system, unoccupied orbitals (virtual orbitals) will be part of the Hartree-Fock solution. Note that, since the virtual orbitals do not contribute to the energy of the system, their shape will not be optimized during the SCF process but they will still be required to be orthogonal to the occupied orbitals. In order to perform an excitation, n occupied orbitals are exchanged with n virtual orbitals. The full configuration interaction (full-CI) wave function contains all possible combinations of excitations of the electrons of the system in the chosen basis set. The coefficients c_i of the expansion of the CI wave function in equation (2.18) are optimized again in an SCF manner while the orbitals conforming the different Slater determinants remain unchanged. This means that the originally virtual orbitals in the Hartree-Fock calculation that were exchanged with occupied orbitals keep their unoptimized shape. In

the limit of a complete basis set, full-CI would give the exact energy but this is again computationally too expensive. For this reason other methods have been developed to approximate the Schrödinger equation.

CIS (CI truncated at the singles) and CISD (CI singles and doubles) recover large parts of the dynamic correlation of the system and are CI methods where n only goes up to 1 or 2, respectively.

If not only the coefficients expanding the Slater determinants c_i in equation (2.18) but also the shape of the orbitals within these Slater determinants $\{\chi_i\}$ are optimized, a multi configuration (MCSCF) calculation is performed. This method usually converges faster than CI and can recover static correlation. CASSCF (complete active space SCF) is a way to perform cheaper MCSCF calculations in which the orbital space and the number of electrons that participate in the excitations are limited [88].

Similar to CI, the wave function of the system can be expanded by excitations from a MCSCF wave function instead of from a HF wave function. A multi reference configuration interaction (MRCI) calculation is then performed [86].

Perturbation theory (PT) [89] is a different approach to improve Hartree-Fock. This method partitions the Hamiltonian into a zeroth-order part $\hat{H}_0$ and a perturbation part $\hat{H}'$:

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}'. \tag{2.19}$$

λ indicates the strength of the perturbation. As λ goes from 0 to 1 the energy and wave function must change continuously. Using a Taylor expansion, they can be written as

$$E = \lambda^0 E_0 + \lambda^1 E_1 + \lambda^2 E_2 + \cdots + \lambda^n E_n + \ldots, \tag{2.20}$$

$$\Psi = \lambda^0 \Psi_0 + \lambda^1 \Psi_1 + \lambda^2 \Psi_2 + \cdots + \lambda^n \Psi_n + \ldots, \tag{2.21}$$

where Ψ_n and E_n are the n-th order corrections to the wave function and to the energy, respectively. When $\lambda = 1$, the n-th order energy or wave function becomes the sum of all terms up to that order. Terms with the same power

of λ can be grouped (as long as the perturbed wave function is intermediately normalized [86]):

$$\lambda^0: \ \hat{H}_0\Psi_0 = E_0\Psi_0, \tag{2.22}$$

$$\lambda^1: \ \hat{H}_0\Psi_1 + \hat{H}'\Psi_0 = E_0\Psi_1 + E_1\Psi_0, \tag{2.23}$$

$$\lambda^2: \ \hat{H}_0\Psi_2 + \hat{H}'\Psi_1 = E_0\Psi_2 + E_1\Psi_1 + E_2\Psi_0, \tag{2.24}$$

$$\lambda^n: \ \hat{H}_0\Psi_n + \hat{H}'\Psi_{n-1} = \sum_{i=0}^{n} E_i\Psi_{n-i}. \tag{2.25}$$

MP2 is the most used (second order) perturbation theory approach. Here $\hat{H}_0$ is the Fock operator and $\hat{H}'$ the difference between the exact Hamiltonian and the Fock energy. Another approach, often used for excited states, is CASPT2 that performs second order perturbation theory from a CASSCF reference energy and wave function.

The last *ab initio* approach that will be discussed here is coupled-cluster theory (CC). CC takes advantage of the fact that the full-CI wave function can be written as [87]:

$$\Psi = e^{\hat{T}}\Psi_{HF}, \tag{2.26}$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_i + \cdots + \hat{T}_N, \tag{2.27}$$

where N is the total number of electrons in the system. $\hat{T}_i$ generates all possible Slater determinants with i excitations from the reference determinant. For example:

$$\hat{T}_1 = \sum_{i}^{occ.}\sum_{a}^{vir.} t_i^a\Psi_i^a, \tag{2.28}$$

$$\hat{T}_2 = \sum_{i}^{occ.}\sum_{j>i}^{occ.}\sum_{a}^{vir.}\sum_{b>a}^{vir.} t_{ij}^{ab}\Psi_{ij}^{ab}, \tag{2.29}$$

where i and j are occupied orbitals while a and b are virtual orbitals, t_i^a and t_{ij}^{ab} are the amplitudes that are determined by the calculation and Ψ_i^a and Ψ_{ij}^{ab} are singly and doubly excited Slater determinants, respectively.

A CCSD calculation, which includes single and double excitations ($\hat{T}_1$ and $\hat{T}_2$), can provide insights on the multireference character of a system via the so-called T1 diagnostic:

$$T_1 = \sqrt{\frac{\sum_{i}^{occ.} \sum_{a}^{vir.} (t_i^a)^2}{N}}. \tag{2.30}$$

A value below 0.02 indicates no multireference character while a value above indicates that multireference methods could be necessary for the correct characterization of the system [128].

CCSD(T), that has become one of the benchmark methods for many systems of interest, explicitly includes single and double excitations and estimates the triple excitations with perturbation theory.

2.1.1.2 Basis Sets

In equation (2.13), the spin orbitals $\{\chi_i\}$ are expanded in a basis set. The basis functions $\{\phi_\mu\}$ could represent atomic orbitals and thus be centered in specific atoms. The most (computationally) convenient basis functions are Gaussian type functions (GTFs) of the form

$$\phi(\mathbf{r}) = (x - X_A)^l (y - Y_A)^m (z - Z_A)^n \exp(-\zeta |\mathbf{r} - \mathbf{R}_A|^2), \tag{2.31}$$

where $\phi(\mathbf{r})$ is the spatial part of the spin orbital centered on nuclei A with coordinates $\mathbf{R}_A = \{X_A, Y_A, Z_A\}$, $\mathbf{r} = \{x, y, z\}$ are the coordinates of a given point in space, ζ is the orbital exponent and l, m, n define the angular momentum quantum number $L = l + m + n$ such that GTFs can be s, p, d, etc. for $L = 0, 1, 2, \ldots$

The use of GTFs is convenient but these functions do not have the correct behavior for hydrogen-like orbitals because they do not have a cusp at the nuclear position and they decay exponentially with r^2 instead of decaying exponentially with r as it should be. For this reason, instead of using a single GTF to approximate a basis function, a linear combination of GTFs (also called "primitive" functions) is used to describe what is called a "contracted" basis function. This also allows atomic orbitals to have a different sign in different parts of space as it happens for the 2s orbital. With this representation, the integrals in equations (2.15) and (2.39) can be analytically evaluated.

Depending on the number of basis functions per occupied atomic orbital, the system can be described by a minimal basis set for one basis function, a double-ζ basis set when using two basis functions per atomic orbital, a triple-ζ basis set for three basis functions and so on. Because the shape of the core orbitals is not usually affected by the formation of bonds, it makes sense to have more flexibility in the valence orbitals than in the core orbitals. This idea is exploited by the so-called split-valence basis sets. Among the most used basis sets of this type are Pople basis sets [90] like for example 6-31G, that is a split-valence-double-ζ basis set with one contracted basis function formed by six primitives for the description of the core orbitals and two basis functions for the description of the valence orbitals one of them with 3 primitives and the other one with just one primitive.

Polarization functions can be added to the basis set to allow for more flexibility. These functions have a larger angular momentum quantum number L. The 6-31G* (also called 6-31G(d)) basis set adds a set of d functions to polarize the p basis functions.

For the description of anions and van der Waals complexes, diffuse basis functions can be used. These functions have small exponents and decay slowly with the distance from the nuclei allowing an electron to be located far from the nuclei. Basis sets that contain this kind of functions are called "augmented" basis sets. In the Pople nomenclature they are denoted with a '+' sign.

Dunning basis sets [91] are another type of commonly used basis sets. They are correlation-consistent which means that the exponents of the primitives were optimized for calculations using methods that include electron correlation (post-Hartree-Fock methods). One of these basis sets is the aug-cc-pVTZ basis set that is used in chapter 7 of this thesis. This is an augmented, correlation-consistent, polarized (with polarization fuctions), valence (meaning it is a split-valence basis set), triple-ζ basis set.

2.1.1.3 Semiempirical Methods*

In the early days of computational chemistry, ways of reducing the computational cost of Hartree-Fock without loosing too much accuracy were desired such that calculations in the computers available at the time would be possible. The most computationally expensive part of Hartree-Fock is the calculation of the Coulomb and exchange integrals (that are two electron integrals) of the Fock operator (equation (2.12)). In order to reduce its cost, only valence electrons were explicitly considered such that the dimensionality of the problem would be reduced. The core electrons were included by modifying the nuclear charges. Moreover, the overlap matrix $\mathbf{S}$ of equation (2.14) was set to the unity matrix so that the Fock matrix $\mathbf{F}$ could be directly diagonalized to give the molecular orbital coefficients and energy levels. Most of the Coulomb and exchange integrals were set to zero. Integrals with values different from zero were parametrized to reproduce experimental data (empirical data) or derived from simple algebraic expressions which enormously reduced the cost of these methods. Of course these approximations come with a cost in the accuracy of the methods. Some methods were further developed and they are a way to get a computationally inexpensive estimation of the behavior of a system although their results should be treated with care. Two of these methods are PM6 [92] and PM7 [93] that are used in chapters 3 and 4 of this thesis, respectively.

2.1.1.4 Density Functional Theory

Ab initio methods are focused on obtaining an approximate representation of the exact electronic wave function from which properties of a system can be calculated. However, the wave function is a complex $(3+1)N$-dimensional object. A simpler object is the electron density [94]:

$$\rho(\mathbf{r}) = N \int \dots \int \Psi^*(\mathbf{x}_1, \dots, \mathbf{x}_N)\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N, \tag{2.32}$$

where N is the number of electrons and $\mathbf{x}_i$ is the spatial coordinate ($\mathbf{r}$) of electron i multiplied by the spin coordinate s. The electron density is a 3-dimensional object that represents the probability of finding any electron at $d\mathbf{r}_1$ independent of

*This section is based on the book by E. G. Lewars *Computaional Chemistry. Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, 2nd ed. (Springer, New York, 2011).

the positions of the other electrons. The electron density gives the total number of electrons when integrated over the whole space. The positions of the nuclei appear as cusps of the electron density and the height of those cusps is determined by the nuclear charges. Thus, the electron density contains enough information to specify the Hamiltonian of the system. Density functional theory (DFT) calculates the properties of a system from the electron density instead of calculating them from a wave function [89]. DFT calculations are based on the Hohenberg-Kohn theorems. The first theorem is an existence theorem that says that any ground state property is a functional of the electron density but does not give an idea about the form of those functionals. The second one is the Hohenberg-Kohn variational theorem that states that any trial electron density will give an energy higher than the exact ground state electron density. Thus, the exact energy of the system could be calculated provided the exact electron density and the form of the energy functional. Unfortunately, neither of them is known. In order to approximate its form, the electronic energy is expressed as a sum of the electronic kinetic energy $\langle T[\rho(\mathbf{r})]\rangle$, the nuclear-electron attraction $\langle V_{\mathrm{Ne}}[\rho(\mathbf{r})]\rangle$ and the electron-electron repulsion $\langle V_{\mathrm{ee}}[\rho(\mathbf{r})]\rangle$:

$$E[\rho(\mathbf{r})] = \langle T[\rho(\mathbf{r})]\rangle + \langle V_{\mathrm{Ne}}[\rho(\mathbf{r})]\rangle + \langle V_{\mathrm{ee}}[\rho(\mathbf{r})]\rangle. \tag{2.33}$$

The problem with equation (2.33) is that the only functional that is known is the nuclear-electron attraction functional. If we imagine a reference system of non-interacting electrons whose electron density is the electron density of the real system, we can define the electronic kinetic energy as

$$\langle T[\rho(\mathbf{r})]\rangle = \langle T[\rho(\mathbf{r})]\rangle_{\mathrm{ref}} + \Delta\langle T[\rho(\mathbf{r})]\rangle, \tag{2.34}$$

were $\langle T[\rho(\mathbf{r})]\rangle_{\mathrm{ref}}$ is the electronic kinetic energy of the reference system and $\Delta\langle T[\rho(\mathbf{r})]\rangle$ the difference between the real electronic kinetic energy and the reference one.

The same can be done for the electron-electron repulsion term that is then defined as

$$\langle V_{\mathrm{ee}}[\rho(\mathbf{r})]\rangle = \langle V_{\mathrm{ee}}[\rho(\mathbf{r})]\rangle_{\mathrm{ref}} + \Delta\langle V_{\mathrm{ee}}[\rho(\mathbf{r})]\rangle. \tag{2.35}$$

The advantage of introducing the definitions in equations (2.34) and (2.35) is that now the only unknowns are $\Delta\langle T[\rho(\mathbf{r})]\rangle$, which represents the kinetic correlation

of the electrons, and $\Delta\langle V_{\text{ee}}[\rho(\mathbf{r})]\rangle$, which represents the exchange and potential energy of the electrons. The sum of these two terms is what is known in DFT as the exchange-correlation energy

$$E_{\text{XC}}[\rho(\mathbf{r})] = \Delta\langle T[\rho(\mathbf{r})]\rangle + \Delta\langle V_{\text{ee}}[\rho(\mathbf{r})]\rangle. \tag{2.36}$$

Combining equations (2.33) to (2.36) the final expression for the DFT energy can be written as

$$E[\rho(\mathbf{r})] = \langle T[\rho(\mathbf{r})]\rangle_{\text{ref}} + \langle V_{\text{Ne}}[\rho(\mathbf{r})]\rangle + \langle V_{\text{ee}}[\rho(\mathbf{r})]\rangle_{\text{ref}} + E_{\text{XC}}[\rho(\mathbf{r})]. \tag{2.37}$$

All terms in equation (2.37) but E_{XC} can be calculated provided an electron density and orbitals (such that the reference electronic kinetic energy can be calculated). Since the reference system has non-interacting electrons, the wave function yielding the reference electron density can be represented as a Slater determinant. In this way the form of the exchange-correlation energy functional is the only unknown. For a Slater determinant wave function the electron density is simply

$$\rho = \sum_i \langle \chi_i | \chi_i \rangle. \tag{2.38}$$

The energy of the system would then be calculated as

$$\begin{aligned} E = \sum_i \left(\left\langle \chi_i \right| -\frac{1}{2}\nabla_i^2 \left| \chi_i \right\rangle - \left\langle \chi_i \right| \sum_A \frac{Z_A}{|\mathbf{r}_i - \mathbf{r}_A|} \left| \chi_i \right\rangle \right) \\ + \sum_i \left\langle \chi_i \right| -\frac{1}{2} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}' \left| \chi_i \right\rangle + E_{\text{XC}}, \end{aligned} \tag{2.39}$$

where $\{\chi_i\}$ are the spin orbitals conforming to the Slater determinant wave function for the reference system, also called Kohn-Sham orbitals. An analogous equation to the Hartree-Fock equation (equation (2.11)) can be written for the calculation of the orbitals that minimize the energy of the system:

$$\hat{h}_i^{\text{KS}} \chi_i = \epsilon_i \chi_i, \tag{2.40}$$

where $\hat{h}_i^{\text{KS}}$ is the Kohn-Sham one-electron operator

$$\hat{h}_i^{\text{KS}} = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{|\mathbf{r}_i - \mathbf{r}_A|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}' + V_{\text{XC}}, \tag{2.41}$$

where V_{XC} is the exchange-correlation potential

$$V_{\mathrm{XC}} = \frac{\partial E_{\mathrm{XC}}[\rho(\mathbf{r})]}{\partial \rho(\mathbf{r})}. \tag{2.42}$$

The exact energy could be calculated given the exact electron density and exchange-correlation energy functional. Since non of this is known, equation (2.40) will be solved in an SCF manner (analogous to the way the Hartree-Fock equation is solved) by expanding the unknown Kohn-Sham orbitals in terms of a basis set. An initial guess for the electron density can be calculated from initial orbitals that will then be minimized with respect to the energy.

The different ways of approximating the form of the exchange-correlation energy functional give rise to the different DFT methods. There are four main ways of approximating this unknown functional form:

1. Local density approximation (LDA): it is the simplest approximation and has an accuracy similar to that of Hartree-Fock. It assumes that the electron density is uniform and thus the expression of the exchange-correlation functional depends only on the electron density.

2. Generalized gradient approximation (GGA): the electron density is not considered as uniform any more. Note that most systems of interest have inhomogeneous electron densities. The exchange-correlation functional depends now on the electron density but also on its first derivative (gradient). Many of them are constructed as a correction term added to an LDA expression. The BLYP functional, for example, is the sum of the B exchange functional, which is based on a correction to an LDA functional, plus the LYP correlation functional that is based on an empirical parametrization of the correlation of the helium atom [95].

3. Meta-GGA: they are an extension of the GGA methods that include the Laplacian (second derivative) of the electron density.

4. Hybrid GGA: they are based on the adiabatic connection [96] that shows that the exchange-correlation energy can be computed as a weighted sum of the DFT exchange-correlation energy and the Hartree-Fock exchange energy. The functionals B3LYP [97] and M06-2X [98] that are used in chapters 3, 4, 5 and 6 of this thesis are functionals of this type that were empirically parametrized.

2.1.2 Force Fields

Force fields (FFs) [89, 95] represent a completely different approach than *ab initio* or DFT methods. They are a mathematical model based on the Born-Oppenheimer approximation. In this model, nuclei are described by balls with a certain mass and charge and are held together by different "springs" modeling the different interactions (see figure 2.1). The potential energy is usually split into bonded and non-bonded contributions:

$$E = E_{\text{bonded}} + E_{\text{non-bonded}}. \tag{2.43}$$

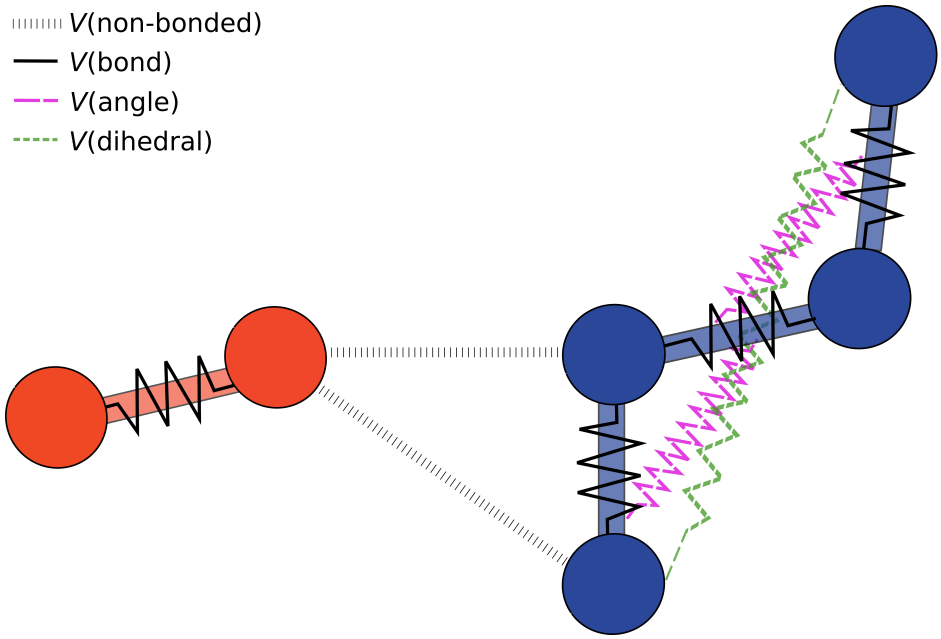


Figure 2.1: Schematic representation of two molecules, one with blue atoms and the other with orange atoms, described by a force field. Balls represent atoms, bars represent bonds, dashed black lines represent non-bonded interactions (V(non-bonded)), black solid lines represent bond potentials (V(bond)), pink broken lines represent angle potentials (V(angle)) and green dashed lines represent dihedral angle potentials (V(dihedral)).

The bonded term is usually a sum of harmonic potentials representing bonds, angles and improper dihedrals plus a periodic potential representing dihedral angles:

$$\begin{aligned} E_{\text{bonded}} = &\sum_{bond} k_r(r - r_0)^2 + \sum_{angle} k_\theta(\theta - \theta_0)^2 \\ &+ \sum_{dihedral} k_\varphi(1 + \cos(n\varphi - \delta)) + \sum_{improper} k_\omega(\omega - \omega_0)^2, \end{aligned} \tag{2.44}$$

where k_r, k_θ and k_ω are the force constants describing bonds, angles and improper dihedrals, respectively. r_0, θ_0 and ω_0 are the equilibrium distances for the harmonic potentials. The third term describes dihedral angles with a periodic potential where k_φ is the force constant, n the periodicity, φ the dihedral angle and δ the phase shift.

The non-bonded term is usually the sum of a Lennard-Jones (LJ) potential and an electrostatic potential:

$$E_{\text{non-bonded}} = \sum_i \sum_{i<j} \left(\epsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right), \tag{2.45}$$

where ϵ_{ij} and $R_{min,ij}$ are the depth and the distance at the minimum of the LJ potential between atoms i and j, r_{ij} is the distance between the two atoms, q_i is the charge of atom i and ϵ_0 is the effective dielectric constant. Note that, in contrast to the methods discussed until now, the bond is a central concept for a FF model and the different interactions are defined from the connectivity of the atoms. One drawback of this description is that most FFs cannot break or form bonds since these are described as harmonic potentials. A way to overcome this limitation is to describe bond terms as Morse potentials that can be dissociated

$$V_{\text{Morse}} = D_e[1 - e^{-\alpha(r-r_0)}]^2, \tag{2.46}$$

where D_e is the depth of the potential well (or the dissociation energy) and α controls the width of the potential. Note that if a bond is actually dissociated, the bonded terms of the FF involving the two atoms that are now disconnected need to be redefined.

Different potentials for the non-bonded interactions can also be used to improve the description of the van der Waals interactions, especially when a bond is

dissociated, since the normal LJ potential can cause unphysically large repulsive forces when the distance between two non-bonded atoms is short. The generalized LJ potential in equation (2.47) has two extra parameters (m, n) that make it more flexible and allow a better description of those regions of the PES where bonds form and break [99]:

$$V_{GLJ} = \frac{n\varepsilon_{ij}}{m-n}\left[\left(\frac{R_{min,ij}}{r_{ij}}\right)^m - \frac{m}{n}\left(\frac{R_{min,ij}}{r_{ij}}\right)^n\right]. \tag{2.47}$$

Some FFs, have additional terms in order to correct the description of the angles, describe hydrogen bonds or account for couplings between the different bonded terms. The CHARMM FF [100] adds Urey-Bradly potentials for the description of angles, which are distance-dependent (instead of angular dependent) harmonic potentials between the outer atoms of an angle [101].

The different potentials modeling the atomic interactions are parametrized to reproduce experimental or reference computational data (an example of a FF parametrization is discussed in section 2.3.3).

2.1.3 Neural Networks

In recent years, machine learning algorithms have started to be used more and more for the description and solution of complex problems. Artificial neural networks are one type of machine learning algorithms that have been proven to be general approximators which makes them well suited to represent a PES. They are inspired on the biological functioning of a brain with its neurons and different connections among them. An artificial neural network has so-called artificial neurons that are connected with each other with a certain architecture that allows the network to "learn" to predict a certain property. High-dimensional neural networks are a type of neural networks that partition the total energy of the system in atomic contributions and use one single neural network for the prediction of the total energy. B. Huang et al. proved that a machine learning model for large molecules can be constructed by training it only on smaller fragments that they called "amons" [102]. This property makes neural networks a promising method to model PESs. The basic transformation in neural networks is a linear regression [103]:

$$\mathbf{y} = \mathbf{W}\ \mathbf{x} + \mathbf{b}, \tag{2.48}$$

where $\mathbf{x}$ is an input vector, $\mathbf{y}$ an output vector and $\mathbf{W}$ and $\mathbf{b}$ are a matrix and a vector of parameters, respectively. In order to approximate an arbitrary function, at least two so-called layers need to be combined

$$\mathbf{y} = \mathbf{W}_2\sigma(\mathbf{W}_1\mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2, \tag{2.49}$$

where σ is an activation function that adds non linearity such that equation (2.49) cannot be collapsed into equation (2.48). In principle, two wide enough layers are sufficient to approximate any relationship, but neural networks composed by more than two layers (deep neural networks) have been proven to be more efficient in their capabilities [103].

In chapter 4 of this thesis, the PhysNet architecture is used. This is a high-dimensional, deep neural network developed by O. Unke et al. [103] that uses nuclear charges and Cartesian coordinates to predict energies, atomic charges, forces and molecular dipoles of chemical systems. The input data for this neural network are nuclear charges and Cartesian coordinates of the system of interest and the "amons" that conform to it. The parameters of the model are optimized to minimize a loss function, which is

$$\begin{aligned}\mathcal{L} = w_E|E - E^{\text{ref}}| + \frac{w_F}{3N}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left|-\frac{\partial E}{\partial r_{i,\alpha}} - F_{i,\alpha}^{\text{ref}}\right| \\ + w_Q\left|\sum_{i=1}^{N} q_i - Q^{\text{ref}}\right| + \frac{w_P}{3}\sum_{\alpha=1}^{3}\left|\sum_{i=1}^{N} q_i r_{i,\alpha} - p_\alpha^{\text{ref}}\right| + \mathcal{L}_{nh}\end{aligned} \tag{2.50}$$

in the present case, and w_E, w_F, w_Q and w_p are the weights of the different contributions to the loss function; E^{ref}, F^{ref}, Q^{ref} and p^{ref} are the reference energies, forces, charges and dipoles, respectively; q_i are partial charges, $\{i\}$ are atoms and $\{\alpha\}$ Cartesian coordinates. $\mathcal{L}_{nh}$ is a "non-hierarchicality penalty" that penalizes situations in which higher-order interactions slowly decay in magnitude since they are known to rapidly decay in many-body expansions. In this way the neural network learns a smooth decomposition of the energy and charge of the system driven only by the training data. The global PES is given by

$$V_{\text{PhysNet}} = \sum_{i=1}^{N} E_i + k_e\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{q_i q_j}{r_{ij}} \tag{2.51}$$

where N is the total number of atoms, k_e is the Coulomb constant, r_{ij} is the distance between atoms i and j, and E_i and q_i are atomic energy contributions and partial charges (corrected to guarantee charge conservation) [103] predicted by PhysNet. Here, the Coulomb's potential is damped at short distances to avoid numerical problems (see Ref. [103]). The PhysNet architecture guarantees that equation 2.51 is invariant with respect to translations, rotations and permutation of atoms sharing the same element type [103].

2.1.4 Optimization of Minima and Transition States

Once we know how to calculate the electronic energy, the PES of a system can be explored. Usually, relevant stationary points are located as a representation of the whole surface. These points are minima that represent stable or metastable states of the system and saddle points (transition states) that connect different stable states. As it was shown at the beginning of this chapter, within the Born-Oppenheimer approximation, the energy of the system is a function of the nuclear coordinates. The problem of exploring a PES is therefore reduced to find those nuclear coordinates that are minima and transition states of the PES.

Mathematically, a stationary point is a point in which all partial derivatives (forces) with respect to each geometric parameter in the vector $\mathbf{q}$ are zero:

$$\frac{\partial V}{\partial \mathbf{q}} = 0. \tag{2.52}$$

A minimum has second order partial derivatives greater than zero (and thus is a minimum in all directions). A first-order saddle point is a minimum in all directions but one that we will call the reaction coordinate and has a negative second order partial derivative. The lowest-energy path connecting two minima through a transition state (TS) is the minimum energy path (MEP) and follows the reaction coordinate. A method often used to calculate the MEP is the intrinsic reaction coordinate (IRC) [104] that calculates the steepest descent path from the TS in mass-weighted Cartesian coordinates. The steepest descent method follows the direction opposite to that with the largest negative gradient until a minimum is reached [95].

There are many algorithms that can be used to find the minima and saddle points of a multidimensional function like the PES. These algorithms vary in

their definition of the direction $\mathbf{q}_0$ along which the nuclear configuration $\mathbf{R}_i$ will be displaced in order to approach a stationary point, and in the size of the step λ used to move the structure along that chosen direction such that

$$\mathbf{R}_{i+1} = \mathbf{R}_i + \lambda \mathbf{q}_0. \tag{2.53}$$

One of the simplest algorithms is "steepest descent" that was briefly discussed above. A more sophisticated method is the Newton-Raphson method [95] that makes a Taylor expansion up to second order around the current configuration $\mathbf{R}_i$. The energy of the next geometry can then be calculated as

$$E(\mathbf{R}_{i+1}) \approx E(\mathbf{R}_i) + \mathbf{g}^t(\mathbf{R}_{i+1})(\mathbf{R}_{i+1} - \mathbf{R}_i) + \frac{1}{2}(\mathbf{R}_{i+1} - \mathbf{R}_i)^t \mathbf{H}(\mathbf{R}_{i+1})(\mathbf{R}_{i+1} - \mathbf{R}_i), \tag{2.54}$$

where $\mathbf{g}^t$ is the transpose of the gradient (the vector with the first partial derivatives of the energy with respect to each internal coordinate) and $\mathbf{H}$ is the Hessian (the matrix with the second order partial derivatives). The condition that the configuration $\mathbf{R}_{i+1}$ is a minimum leads to

$$(\mathbf{R}_{i+1} - \mathbf{R}_i) = -\mathbf{H}^{-1}\mathbf{g}. \tag{2.55}$$

The adopted basis Newton-Raphson (ABNR) method used in chapters 4 and 5 of this thesis is a mixture between the steepest descent and the Newton-Raphson methods [105].

Finding transition states is more challenging than finding minima since one of the directions must be maximized. One of the algorithms that can deal with such a problem is conjugate peak refinement [106] that finds the saddle point(s) connecting two given minima. It is based on the idea that, in the vicinity of a saddle point, there will be a local maximum along one of the directions of the PES. In this method the energy is again approximated as quadratic like in equation (2.54). The first guess for the direction of the reaction coordinate is simply the difference of the coordinates of the two minima. This direction is iteratively refined until the saddle point is reached.

Most optimization algorithms need a first guess of the structure and can only find the closest minimum to that initial guess. However, there are some programs that, starting from a minimum, can make an automatized exploration of the

PES. GRRM (global reaction route mapping) [107] is one of these algorithms. It is based on the anharmonic downward distortion following (ADDF) method [108]. The idea behind ADDF is that, if the minima could be described by harmonic wells, dissociation potentials and neighboring minima would distort the harmonic form of a well in the direction of the reaction coordinate. By following uphill this anharmonically distorted direction, a TS can be found. From a TS an IRC can be calculated to get the new minimum connected to the initial minimum and from there find other TSs and other minima.

2.2 Kinetic Studies*

An important aspect once the topology of a PES has been characterized is the speed at which the system moves from minimum to minimum by crossing the different barriers on that PES. Transition state theory (TST) was one of the first theories to connect the notion of a reaction coordinate with the rate at which the system moves along this coordinate [109]. A simplified PES containing only two minima connected by a TS is depicted in figure 2.2. The potential energy of the TS structure determines the height of the barrier ΔE_0 that the reactants (R) would have to overcome to bring the system to the product state (P).

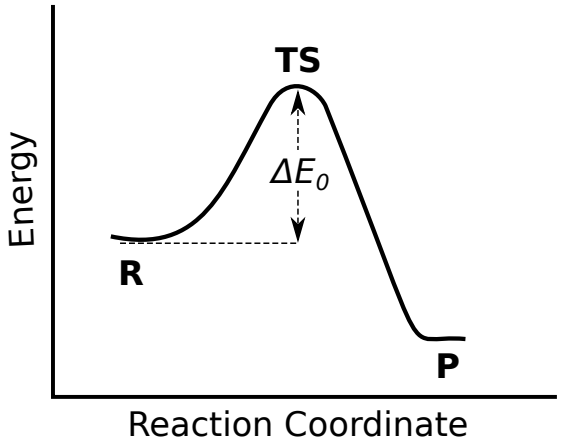


Figure 2.2: Schematic representation of a simplified potential energy surface along the reaction coordinate. The reactant (R) and product (P) states as well as the transition state (TS) are labeled. ΔE_0 is the energy difference between the reactants and the transition state.

*This section is based on the book by R. D. Levine *Molecular Reaction Dynamics* (Cambridge University Press, Cambridge, 2005).

Systems with kinetic energy rarely visit the TS structure but an ensemble of structures that resemble the TS structure. TST defines different regions on the potential energy surface: the reactant region, the product region and a hypersurface of $3N-7$ dimensions that separates them. In TST, if the system crosses the dividing hypersurface, it will go to products without ever crossing back to reactants. In a system with total energy E, the excess energy at the dividing hypersurface may be partitioned into the energy in the reaction coordinate ε_{rc} and the energy in the other internal degrees of freedom of the reactant ε_{int}:

$$E - \Delta E_0 = \varepsilon_{rc} + \varepsilon_{int}. \tag{2.56}$$

In this way, the reaction coordinate is treated classically as a translation and assumed to be separable from the perpendicular directions. No quantum effects are taken into account (although different methods exist to include tunneling [110, 111]). A TS-like structure with energy ε_{rc} in the reaction coordinate will have a linear momentum p_{rc} and velocity v_{rc} along the reaction coordinate such that

$$p_{rc} = m_{rc} v_{rc} = \sqrt{2 m_{rc} \varepsilon_{rc}}, \tag{2.57}$$

where m_{rc} is the mass for the unidimensional translation along the reaction coordinate. The rate of barrier crossing in the energy range $[\varepsilon_{rc},\ \varepsilon_{rc} + d\varepsilon_{rc}]$ will be

$$dr_{rc} = v_{rc}\ dN_{rc}, \tag{2.58}$$

where dN_{rc} is the number of systems along one unit length of the reaction coordinate. According to quantum mechanics this is

$$dN_{rc} = \frac{dp_{rc}}{h}, \tag{2.59}$$

where h is the Plank's constant. Taking equation (2.59) into account, equation (2.58) can be rewritten as

$$dr_{rc} = v_{rc} \frac{dp_{rc}}{h} = \frac{d\varepsilon_{rc}}{h} \tag{2.60}$$

which indicates that the rate of barrier crossing is only determined by the energy in the reaction coordinate at the dividing hypersurface and it is independent of how the system got to that point. However, the excess energy may be partitioned in different ways among the internal degrees of freedom of the system. Since TST

assumes the system in the reactant state to be in equilibrium, all energy partitions are equally probable. Thus, the total rate of barrier crossing will be the sum of all possible states with an energy $d\varepsilon_{rc}$ in the reaction coordinate:

$$\sum_i dr_{rc} = \sum_i \frac{d\varepsilon_{rc}}{h} \equiv \frac{dE}{h} N^{\ddagger}(E - \Delta E_0), \tag{2.61}$$

where i runs over the internal degrees of freedom, $d\varepsilon_{rc}$ is assumed to be the same quantity as dE and $N^{\ddagger}(E - \Delta E_0)$ is the number of states of the system at the dividing hypersurface taking into account all internal degrees of freedom but the reaction coordinate. The cumulative rate per unit energy $Y(E)$ can then be written as

$$Y(E) = \frac{1}{h} N^{\ddagger}(E - \Delta E_0). \tag{2.62}$$

The rate constant can be calculated dividing the cumulative rate by the concentration of reactants. In this case, the concentration is the number of reactant states per unit volume and unit energy $\rho(E)$ (what we call density of states):

$$k(E) = \frac{N^{\ddagger}(E - \Delta E_0)}{h\rho(E)}. \tag{2.63}$$

RRKM (Rice-Ramsperger-Kassel-Marcus) theory [112] takes TST one step further for unimolecular reactions allowing to study systems with more than one TS along the reaction coordinate. Its basic assumption is the ergodicity assumption that states that the energy is redistributed among the degrees of freedom of the molecule (normal modes) faster than the velocity at which the system can cross the different barriers. In this way the individual processes may be treated independently from each other. TST does not take into account how the reactant system gets to the dividing surface. On the contrary, in RRKM theory the reactant system (R) is in equilibrium with an energized species (R*) that can cross the dividing hypersurface to form products

$$\mathrm{R} \rightleftharpoons \mathrm{R}^* \rightarrow \mathrm{P}. \tag{2.64}$$

If the energized species live long enough, their internal states will be in equilibrium and TST may be used. With these assumptions, the rate constant in equation (2.63) can be obtained where now $\rho(E)$ will be the density of states of the energized species.

In order to calculate $N^{\ddagger}(E-\Delta E_0)$ and $\rho(E)$, the molecules are thought of as sets of harmonic oscillators such that

$$N(E) = \frac{1}{s!}\prod_{i=1}^{s}\frac{E}{h\nu_i}, \tag{2.65}$$

$$\rho(E) = \frac{dN(E)}{dE}, \tag{2.66}$$

where s is the number of harmonic oscillators and ν_i the frequency of oscillator i. More sophisticated ways of counting states, where anharmonic effects and rotational states are also taken into account, have been developed [113].

Once the rates of barrier crossing have been calculated, they can be used to follow the kinetics of a system across a PES by solving a master equation of the form

$$\frac{dP_i(t)}{dt} = -\sum_{j\neq i} k_{ij}(E)P_i(t) + \sum_{l\neq i} k_{li}(E)P_l(t), \tag{2.67}$$

where P_i and P_l are the population of states i and l at time t. $k_{ij}(E)$ and $k_{li}(E)$ are the RRKM rates to go from state i to state j and from state l to state i, respectively, at an energy E.

2.3 Molecular Dynamics Simulations

The phase space is the $6N$-dimensional space that contains all possible states of a system constituted by N particles. Each of those particles has 3 spatial coordinates and 3 momentum coordinates. In classical molecular dynamics, Newton's second law of motion ($F = ma$, where F is the force, m is the mass and a the acceleration) is used to follow the system across the phase space as a function of time. A point in phase space is defined by the positions and momenta of the particles constituting that system which determine the potential and kinetic energy at that point. Thus, a point in phase space already determines the next point that the system will visit. In practice, the next point is calculated using a propagation method [95].

For a given PES, the forces acting on each particle of the system $\mathbf{F}_i$ can be calculated as

$$\mathbf{F}_i = \frac{\partial V}{\partial \mathbf{q}_i}, \tag{2.68}$$

where $\mathbf{q}_i$ are the 3 spatial coordinates of particle i and V is the potential energy calculated with one of the methods described in section 2.1. The relationship between the position of a particle at time t_1 and the position of that particle at time t_2 is simply

$$\mathbf{q}_i(t_2) = \mathbf{q}_i(t_1) + \int_{t_1}^{t_2} \frac{\mathbf{p}_i(t)}{m_i} dt = \mathbf{q}_i(t_1) + \int_{t_1}^{t_2} \mathbf{v}_i(t) dt, \tag{2.69}$$

where $\mathbf{p}_i(t)$ and $\mathbf{v}_i(t)$ are the momentum and velocity of particle i at time t and m_i is the mass of that particle. Since the integral in equation (2.69) is not trivial to solve for most systems, Euler's approximation could be used such that

$$\mathbf{q}_i(t + \Delta t) = \mathbf{q}_i(t) + \mathbf{v}_i(t)\Delta t, \tag{2.70}$$

where the velocity is calculated as

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \mathbf{a}_i(t)\Delta t \tag{2.71}$$

with Δt being the time step and $\mathbf{a}_i(t)$ the acceleration of particle i at time t that can be calculated according to Newton's second law of motion from the forces in equation (2.68). In the limit that $\Delta t \to 0$, equations (2.70) and (2.71) are exact. However, an infinitesimally small time step would make the calculation of a trajectory unfeasible due to its huge computational cost. The size of Δt determines for how long the system will move in a certain direction (with a certain velocity) until new velocities are calculated. A large time step would be computationally cheaper but could lead to a situation in which two particles are too close to each other. This situation would result in high repulsive energies that would cause chemical instabilities. Generally, Δt is chosen to be smaller than the fastest periodic motion in the system. In most chemical systems of interest, the fastest motions are bond stretchings concerning hydrogen atoms with a period of $\approx 10^{-14}$ s and thus a typical time step is $\Delta t = 0.1$ fs.

Even with a small time step, the integration scheme in equations (2.70) and (2.71) is not as stable as it would be desired. More complex and stable integration schemes have been developed. One of them is the velocity Verlet algorithm that

is used in this thesis [114]. In this algorithm, coordinates and velocities are propagated as follows:

$$\mathbf{q}_i(t+\Delta t) = \mathbf{q}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{1}{2}\mathbf{a}_i(t)\Delta t^2, \tag{2.72}$$

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + \frac{1}{2}\left[\mathbf{a}_i(t) + \mathbf{a}_i(t+\Delta t)\right]\Delta t. \tag{2.73}$$

The main advantage of this algorithm is that positions and velocities are calculated at the same time (in the leapfrog algorithm [115], an alternative propagation method, this is not the case). Moreover, velocity Verlet is time-reversible and conserves total energy.

A trajectory can be recorded by propagating initial coordinates and velocities using equations (2.72) and (2.73) in an iterative way.

2.3.1 Preparation of Initial Conditions

As mentioned above, initial coordinates and velocities must be provided in order to propagate Newton's equations of motion. The initial conditions are usually drawn from statistical ensembles in order to get meaningful trajectories. A statistical ensemble is a large number of replicas of the system in different points of phase space (in different states) that share some common macroscopic properties. There are different kinds of ensembles depending on the magnitudes chosen to define the ensembles. Two examples are the microcanonical or NVE ensemble, with constant number of particles N, volume V and total energy E, and the canonical or NVT ensemble, with constant number of particles N, volume V and temperature T. An NVE ensemble can be generated in the following way:

1. The geometry of the system (defined by N particles and volume V) is minimized. This step initializes the coordinates $\mathbf{q}$.

2. Initial velocities $\mathbf{v}$ are initialized from a Boltzmann distribution at a certain low temperature T_1.

3. Heating dynamics: according to the equipartition theorem, the temperature of an ensemble can be calculated from the average kinetic energy of the ensemble (see equation (2.74)). Thus, the temperature is proportional to the atomic velocities and it can be raised from a low temperature T_1 to

the final temperature T by scaling up the velocities for a certain number of simulation steps.

$$T = \frac{2}{3Nk_B} \left\langle \sum_i^N \frac{1}{2} m_i \mathbf{v}_i^2 \right\rangle \tag{2.74}$$

4. Equilibration dynamics: the temperature is kept constant at a value T by rescaling the velocities at regular intervals.

5. Production dynamics: the ensemble evolves freely. Since the velocities are no longer monitored to keep a certain temperature, this is now an NVE ensemble.

In practice, steps 1 to 5 can be performed only once with a long production dynamics. Snapshots (velocities and coordinates) at different time steps of the production dynamics can be extracted and propagated further to ensure that the initial conditions that will be used for the calculation of ensemble properties define different points in phase space.

When the system concerns a single molecule, most of the kinetic energy will be found in the form of vibrational energy. For simulations of collision experiments, collision energy can be introduced by scaling the atomic velocities along the collision axis and the impact parameter can be varied by displacing the reactant molecules along a perpendicular direction.

Rotational energy can be included in the following way: the moment of inertia tensor of the molecule can be calculated and from its diagonalization, the axis of inertia and the rotational constants of the molecule. When the molecule is aligned with its principal axis of inertia, according to the equipartition theorem, rotational energy E_{rot} can be added as:

$$E_{\text{rot},\alpha} = \frac{1}{2} k_B T = \frac{1}{2} I_\alpha w_\alpha^2, \tag{2.75}$$

where $\alpha = \{x, y, z\}$ and I_α and w_α are the rotational constant and the angular velocity in the α direction, respectively. Reorganizing the previous equation, the angular velocity corresponding to a certain temperature T can be calculated as

$$w_\alpha = \sqrt{\frac{k_B T}{I_\alpha}}. \tag{2.76}$$

The atomic linear velocity, calculated as the cross product between the atomic angular velocity and the atomic coordinates (equation (2.77)), can be added to the atomic velocities of the ensemble:

$$\mathbf{v}_i = \mathbf{w}_i \times \mathbf{q}_i. \tag{2.77}$$

Excitation of specific vibrational normal modes can be performed by projecting the velocities onto the eigenvectors of the Hessian matrix and modifying the kinetic energy of the desired mode. The modified velocities can then be projected back to the Cartesian space.

2.3.2 Analysis of Trajectories

Once an ensemble of trajectories has been computed, different data can be extracted from it. One of the most important objectives is the identification of reactive trajectories. Reactivity must be understood here as a broad concept. What we want to know is whether or not the process of study took place over the course of the trajectory. This condition can be defined in terms of the internal coordinates of the system. For example, in a collision experiment, if at the end of the trajectory the distance between the centers of mass of the colliding molecules is shorter than a threshold, a van der Waals complex has been formed. If a reaction is being studied, the length of the bonds to be formed might be a good choice for the description of reactants and products. In this way, if this length is always larger than a threshold we know that the trajectory does not sample a reactive event. Whereas when this length is shorter than the defined threshold for a certain number of steps in the trajectory, a reaction has taken place.

In simulations of collision experiments, cross sections and reaction rates can be calculated provided a sufficiently broad sampling of initial conditions [116]. For a uniform sampling of the impact parameter b, the cross section σ can be calculated as

$$\sigma = 2\pi b_{\max} \frac{1}{N_{\text{tot}}} \sum_{i=1}^{N_{\text{reac}}} b_i, \tag{2.78}$$

where $b_{\max}$ is the maximum impact parameter defined as the impact parameter at which no reactions can be observed anymore, N_{tot} is the total number of

trajectories, N_{reac} is the number of reactive trajectories and b_i is the impact parameter of the reactive trajectory i.

A rate of reaction can be calculated simply by multiplying the reactive cross sections by the relative velocities of the molecules involved in the collision.

2.3.3 Multi-Surface Adiabatic Reactive Molecular Dynamics

Most of the conventional FFs cannot be used for the study of reactions which involve bond breaking and forming. Multi-surface adiabatic reactive molecular dynamics (MS-ARMD) [99] is a method that circumvents this limitation. In MS-ARMD, the different connectivities of the system (that result from forming and/or braking bonds) are described by different FFs. These representations are then connected with smoothing functions. The global PES is described by

$$V_{MS-ARMD} = \sum_{i=1}^{n} w_i(\mathbf{x}) V_i(\mathbf{x}) + \sum_{i=1}^{n-1} \sum_{j=1+1}^{n} [w_i(\mathbf{x}) + w_j(\mathbf{x})] \sum_{k=1}^{n_{ij}} \Delta V_{GAPO,k}^{ij}(\mathbf{x}), \quad (2.79)$$

where $V_i(\mathbf{x})$ is the energy of the FF i at the nuclear geometry $\mathbf{x}$, $\Delta V_{GAPO,k}^{ij}(\mathbf{x})$ are Gaussian times polynomial (GAPO) functions used to smooth the crossing region between the different FF representations with the form

$$\Delta V_{GAPO,k}^{ij}(\mathbf{x}) = exp\left(-\frac{(\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^{0})^2}{2\sigma_{ij,k}^2}\right) \sum_{l=0}^{m_{ij.k}} a_{ij,kl} (\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^{0})^l, \quad (2.80)$$

where $\Delta V_{ij}(\mathbf{x})$ is the difference in energy between the FFs V_i and V_j at the nuclear configuration $\mathbf{x}$, $V_{ij,k}^{0}$ and $\sigma_{ij,k}$ are the center and the width of the k-th Gaussian function and $a_{ij,kl}$ is the l-th coefficient of the polynomial function of order $m_{ij.k}$ that defines the k-th GAPO function. $w_i(\mathbf{x})$ is the weight of $V_i(\mathbf{x})$ with the form

$$w_i(\mathbf{x}) = \frac{w_{i,0}(\mathbf{x})}{\sum_{j=1}^{n} w_{j,0}(\mathbf{x})}; \quad w_{i,0}(\mathbf{x}) = exp\left(-\frac{V_i(\mathbf{x}) - V_{min}(\mathbf{x})}{\Delta V}\right), \quad (2.81)$$

where $V_{min}(\mathbf{x})$ is the energy for the FF that gives the lowest value at the configuration $\mathbf{x}$ and ΔV is a parameter.

In order to get the different FF representations $V_i(\mathbf{x})$, an initial parametrization is used for the description of each of the states of the system. For the example

in figure 2.2, a FF for the reactant state (R) and a FF for the product state (P) are generated. The initial parametrization is iteratively refined by fitting the parameters of the FFs to reference energies with a simplex algorithm [117]. The new parametrization is used to generate new structures (see section 2.3.1) whose energies will be added to the reference energies and used for further refinement of the FF parameters. The parametrization finishes when the FF correctly describes the energies of newly generated structures.

The downhill simplex algorithm is a general optimization method in which a polyhedron of $N + 1$ dimensions samples an N-dimensional function to find a minimum. This method replaces the vertex of the simplex with the highest value of the function P_h in an iterative way by a new point P^*. The new point is calculated by a reflection operation where this point is reflected along the line that joins it to the centroid of the simplex $\bar{P}$. If the new point P^* has now the lowest value of the function, an expansion operation is performed, meaning that the newly calculated point is moved farther along that direction. If on the contrary, the newly generated point still has the highest value of the function, a contraction is performed where the point is displaced into the negative direction. Finally, if this new point P^* has a value of the function that is neither the highest nor the lowest, the point is accepted and the current vertex with the highest value is selected by the algorithm to continue the process.

When the two FFs are constructed, the GAPO functions can be fitted to reproduce the IRC of the reaction calculated with the reference method as shown in figure 2.3. At the beginning of the reaction coordinate, the system is described by the reactant FF while at the end of the reaction coordinate, it is described by the product FF. In the intermediate region, the mixing of the two FFs with the smoothing GAPO functions takes place to reproduce the IRC of the model reaction. The GAPO functions are fitted using a genetic algorithm [118, 119]. This algorithm is based on the idea of natural selection such that the fittest (best) solutions are those that survive until the algorithm converges. Starting from an initial population of individuals (possible solutions) that are represented by genes (parameters or variables), a selection based on a fitness function is performed to choose the individuals that are more promising. New generations (iterations) are created then from the previously chosen individuals by two different genetic operations, namely mutation and crossover. In the crossover a new possible solution is calculated from the mixture of the genes of two individuals while in the mu-

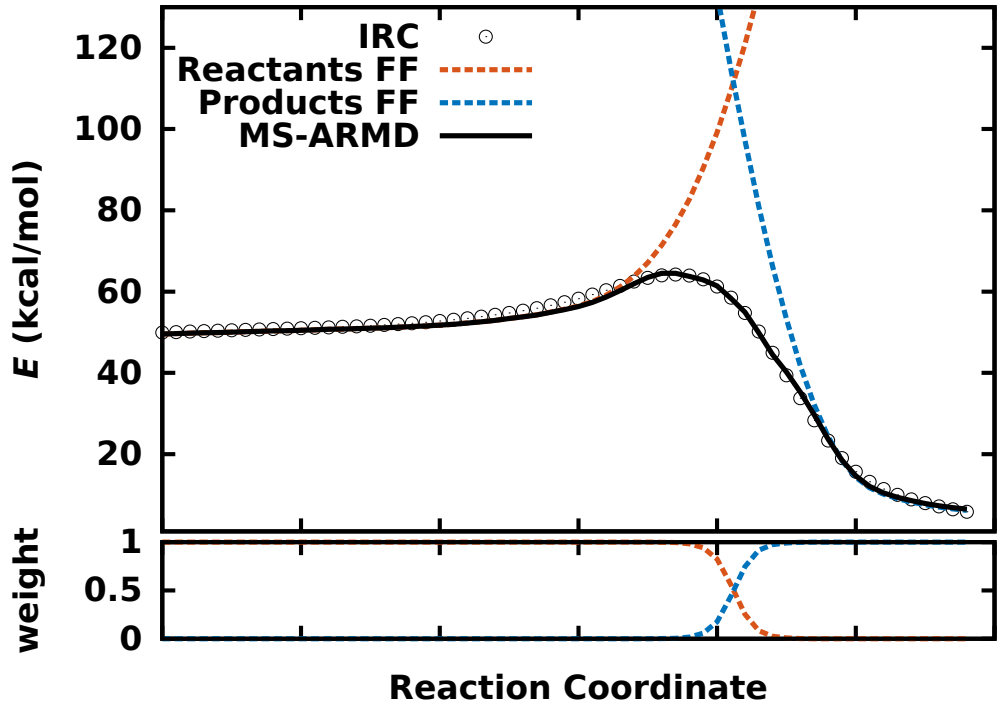


Figure 2.3: Example of multisurface adiabatic reactive molecular dynamics potential energy surface (in black) constructed from two force fields that describe the reactants and products of a model system in orange and blue, respectively. The reference energies are represented by dots. The bottom panel shows the weights of the force fields along the reaction coordinate.

tation some of the genes are randomly changed with a certain probability. The algorithm finishes when the offspring is not better than the parent generation.

Chapter 3

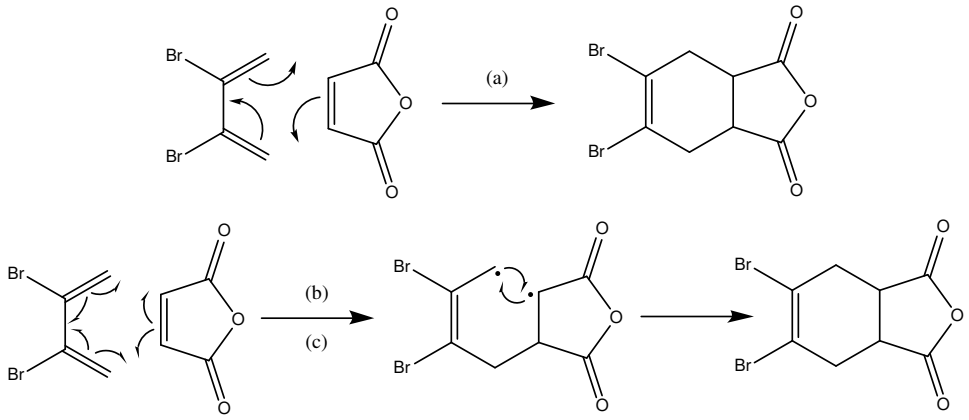
Energetics and kinetics of DBB + MA *

3.1 Introduction

In this chapter, the neutral and cationic gas-phase Diels-Alder reactions between 2,3-dibromo-1,3-butadiene (DBB) and maleic anhydride (MA) (see scheme 3.1) are computationally explored as the first step of a combined experimental and theoretical study. By studying both variants at the same time we can explore the potential differences between neutral and cationic mechanisms discussed in section 1.2 such as the (a)symmetry and (a)synchronicity of the processes.

The vast majority of experimental data is available for solution phase. A crucial difference when studying reactions in the gas phase under single-collision conditions, is that the energy released into the products over the course of the reaction cannot be quenched by the environment on the timescales of experiments and is therefore available to drive consecutive fragmentations or isomerizations. This effect is particularly important for Diels-Alder processes in which the total energy release of the reaction remains locked in a single product. As only scarce information is available on these important mechanistic aspects [120], a comprehensive theoretical characterization of the mechanisms, decay pathways and kinetics of

*This chapter is based on the paper *A computational study of the Diels-Alder reactions between 2,3-dibromo-1,3-butadiene and maleic anhydride.* U. Rivero, M. Meuwly, S. Willitsch, Chemical Physics Letters 683 (2017) 598.



Scheme 3.1: Diels-Alder reaction between 2,3-dibromo-1,3-butadiene (DBB) and maleic anhydride (MA): (a) Concerted mechanism, (b) and (c) stepwise mechanisms with a short-lived and a long-lived intermediate, respectively.

the products of these reactions are presented here which highlights these effects and may serve as a guide to future experiments.

As discussed in section 1.3, we have chosen DBB as the diene because it is a generic, activated diene which fulfills the experimental requirements for conformational separation of its isomers by electrostatic deflection of a molecular beam [26] thus enabling the characterization of conformational aspects and specificities of the reaction. MA is a widely used, activated dienophile which due to its symmetry simplifies the possible outcomes of the reaction. The reaction of DBB and MA thus serves as a prototypical system well suited to explore general mechanistic aspects of Diels-Alder processes under gas-phase conditions.

Computational results on this system using DFT and multi-reference (CASPT2) approaches are presented in Section 3.3.1. The different isomerization and fragmentation pathways that the products of the Diels-Alder reactions can follow are shown in Section 3.3.2. Finally, in order to obtain qualitative trends for the time evolution of the Diels-Alder products under collisionless conditions, a study based on RRKM theory [112] is performed in Section 3.3.3.

3.2 Methods

The PES of the system was calculated at the DFT level of theory with the M06-2X functional [98] as recommended by Linder and Brinck [63] and using the Gaussian09 suite of codes [121]. The 6-31G* basis set was used as recommended in

Refs. [62, 68]. We have employed the GRRM14 program [122] to scan the PES for stationary points by performing a global reaction route mapping (GRRM) calculation [107] at the PM6 level of theory [92] to investigate the isomerization and fragmentation pathways of the products of the Diels-Alder reactions. The stationary points found both by GRRM and by manual exploration of the PES have been optimized at the B3LYP/6-31G* level [97] and then reoptimized at the M06-2X/6-31G* level of theory. All the geometry optimizations have been followed by frequency calculations and the connectivities of the TSs with the minima have been checked by means of internal reaction coordinate (IRC) calculations [104]. The energies have been corrected for zero-point vibrational energies without scaling the frequencies. Single point calculations of some of the stationary points at the CASPT2(5,5)/aug-cc-pVDZ level of theory as implemented in Molcas 8.0 [123] were performed.

We have employed RRKM theory in combination with a master equation approach (equation (3.1)) to follow the time evolution of the products of the Diels-Alder reactions until their eventual fragmentation [112].

$$-\frac{dP_i(t)}{dt} = \sum_{j \neq i} k_{ij}(E) P_i(t) - \sum_{l \neq i} k_{li}(E) P_l(t) \tag{3.1}$$

$P_i(t)$ is the population of species i at time t and $k_{ij}(E)$ is the microcanonical RRKM rate for the formation of species j from species i at energy E. The different $k_{ij}(E)$ have been calculated using Multiwell [124]. Tunneling has been taken into account using an asymmetric Eckart barrier model [110]. We have used Maple 2015 [125] with the Rosenbrock numerical method for solving equation (3.1) [126].

3.3 Results

3.3.1 Diels-Alder reaction

Figures 3.1 and 3.2 show stationary points along the reaction coordinate for the neutral and cationic Diels-Alder reactions between DBB and MA at the M06-2X/6-31G* and at the B3LYP/6-31G* levels of theory. Because both reactant molecules are symmetric, there are two possible paths for the Diels-Alder reaction to follow. These are referred to as “endo” and “exo” depending on the relative

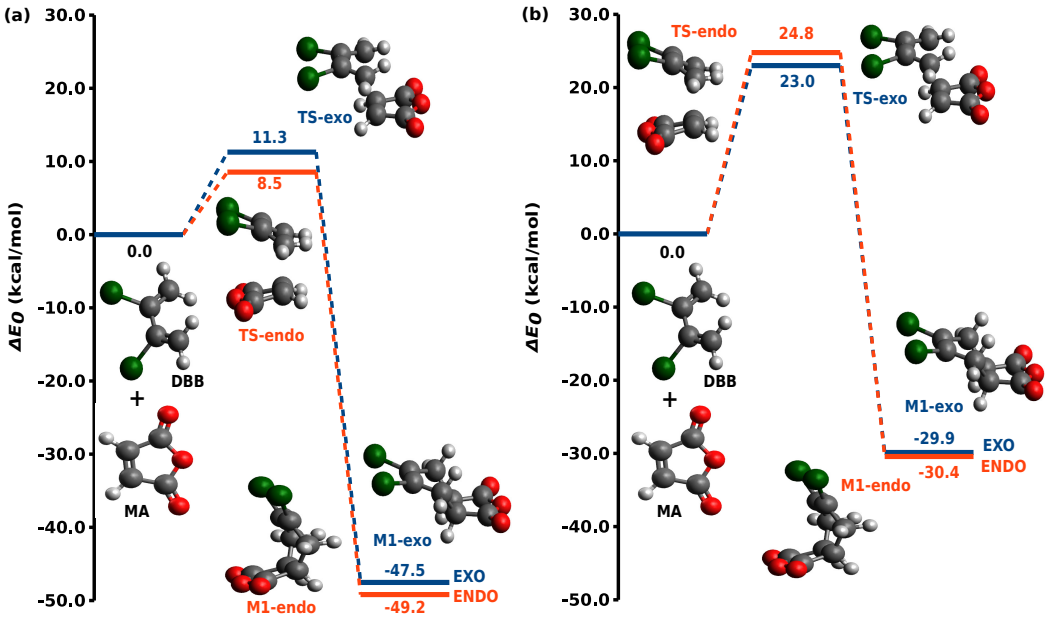


Figure 3.1: Potential energy surface for the two possible neutral Diels-Alder reactions (exo in blue and endo in orange) between s-*cis*-2,3-dibromo-1,3-butadiene (DBB) and maleic anhydride (MA) at the (a) M06-2X/6-31G* and (b) B3LYP/6-31G* levels of theory. The relative energies in kcal/mol with respect to the reactants as well as the structures of minima and transition states are shown. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

orientation of the reactants. M06-2X clearly favors the endo path over the exo as expected for this kind of reactions [127]. On the contrary, B3LYP favors the endo path and has a lower enthalpy of reaction than M06-2X.

Judging from the PESs and the geometries of the TSs, the neutral reaction (figure 3.1) is concerted and symmetric. This is not a surprise since, even though s-*cis*-DBB is not planar, it becomes planar when it interacts with MA making the TSs symmetric.

The situation for the cationic case is quite different (figure 3.2). With an odd number of π electrons, the symmetry is broken and this is no longer a [4+2] but a [3+2] cycloaddition so the TSs are now asymmetric. Interestingly, the endo path is predicted to be concerted while the exo path is so asymmetric that it switches to a stepwise mechanism with a shallow intermediate between the two TSs. Because a stepwise mechanism becomes favorable, the s-*trans* conformer of DBB can also participate in the reaction. We were able to locate one path (in green in figure 3.2) connecting the reaction of the s-*trans* conformer with the intermediate of the exo path. Again B3LYP gives a lower enthalpy of reaction

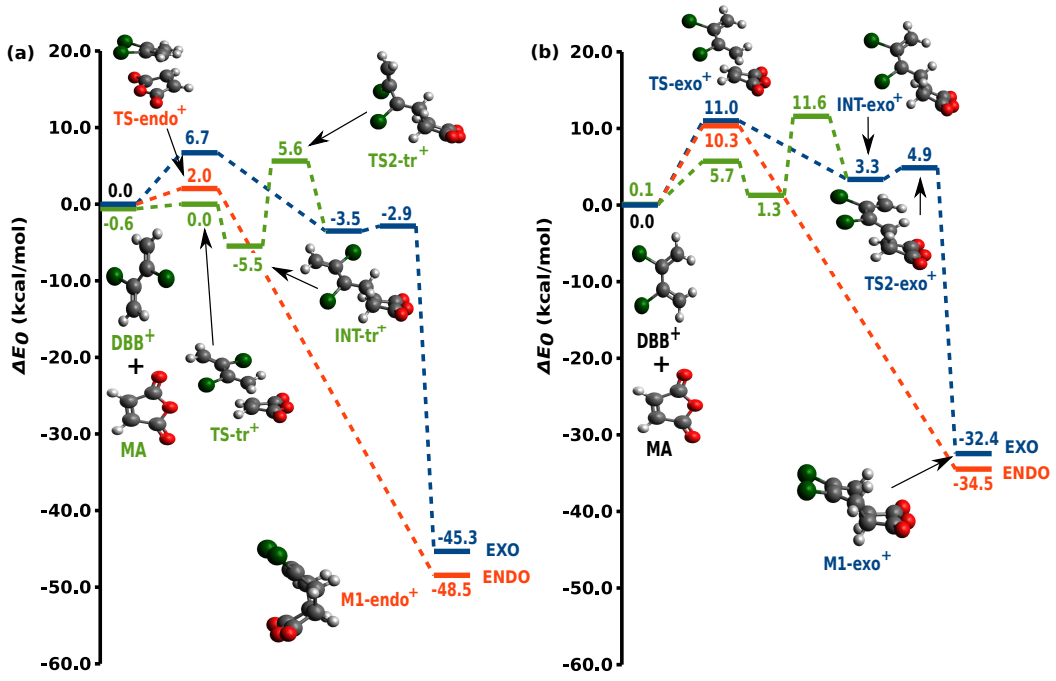


Figure 3.2: Potential energy surface for the three possible cationic Diels-Alder reactions (exo in blue, endo in orange and with the s-*trans* conformer in green) between cationic 2,3-dibromo-1,3-butadiene (DBB^+) and maleic anhydride (MA). Zero point corrected relative energies in kcal/mol with respect to the reactants at infinite distance at (a) the M06-2X/6-31G* and (b) B3LYP/6-31G* levels of theory. The structures of minima and transition states are shown. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

than M06-2X. In order to explore the multireference character of the cationic surface, single point calculations at the CASPT2(5,5)/aug-cc-pVDZ level of theory were carried out (see figure 3.3). Note that no zero point energy is being considered and that the zero of energy corresponds to the two reactant molecules (with DBB in its s-*cis* conformation) being 7 Å apart instead of at infinite distance. It can be seen that CASPT2 largely agrees with M06-2X but the height of the barriers is reduced. We have also performed single point calculations at the MS(2)-CASPT2(5,5)/aug-cc-pVDZ level of theory finding that the energetic difference between the ground state and the first excited state surface is around 25 kcal/mol or higher at all the stationary points of figure 3.3. This is also in line with a CCSD/aug-cc-pVDZ single point calculation that we performed for M1-endo (since it is the stationary point where the ground state surface and the first excited state surface approach each other the most) obtaining a value of the T1 diagnostic of 0.0207. It is usually argued that values below 0.02 indicate that a single reference wave function is sufficient to treat a system while values

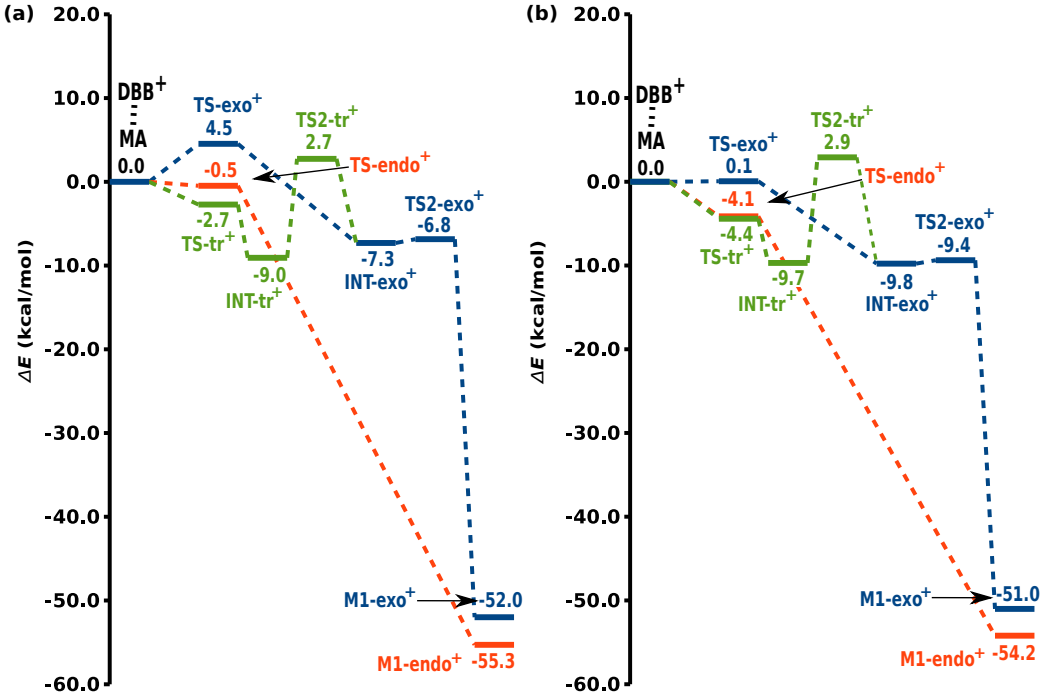


Figure 3.3: Potential energy surface for the three possible cationic Diels-Alder reactions (exo in blue, endo in orange and with the s-*trans* conformer in green) between cationic 2,3-dibromo-1,3-butadiene (DBB^+) and maleic anhydride (MA). Relative electronic energies in kcal/mol with respect to the reactants (with DBB^+ in its s-*cis* conformation) at a distance of 7 Å at (a) the M06-2X/6-31G* and (b) CASPT2/aug-cc-pVDZ levels of theory respectively.

above 0.02 indicate multireference character [128]. We therefore believe that the single-reference M06-2X/6-31G* method should be sufficient for our purposes, i.e., studying the qualitative aspects of the reactivity of this system. It is worth to realize that, in the exo path, the closure of the ring is predicted to have a lower activation barrier (0.6 kcal/mol) than the isomerization into the s-*trans* intermediate (9.1 kcal/mol). Thus, a stereo-selective reaction should be favored even if the mechanism is stepwise. Further dynamic (and experimental) studies are needed in order to confirm this hypothesis. The activation barriers in the cationic reaction (2.0 kcal/mol and 6.7 kcal/mol) are much lower than those in the neutral system (11.3 kcal/mol and 8.5 kcal/mol) implying faster kinetics for the ionic variant.

3.3.2 Isomerizations and fragmentations of the Diels-Alder product

We are interested in the possible isomerizations and fragmentations of the Diels-Alder products to explore their reactivity under typical gas-phase single-collision conditions. Subsequent reactions which have been identified are shown in figures 3.4 and 3.5. Note that only one backward reaction was considered because, with the energy available in the system after the formation of the Diels-Alder product, the endo and exo product species can freely interconvert. Additionally, in the construction of the energy profiles, we only considered structures with different atom connectivities to be different minima. For this reason, rotational and conformational isomers are not distinguished here. We chose the most stable structure among the ones we found to represent each family of isomers. Finally, we only show the most favorable dissociation from each minimum.

The product of the neutral Diels-Alder reaction (M1) can undergo a dissociation into CO + CO_2 + P1 via TS1-P1 or two kinds of isomerizations (figure 3.4): (i) in blue, migration of a hydrogen atom via TS1-2 and TS2-3 followed by migration of bromine via TS3-4 and another migration of hydrogen via TS4-5 or (ii) in orange, migration of a hydrogen atom that breaks the 5-membered ring via TS1-6, followed by another hydrogen migration (TS6-7) and an -OH migration (TS7-8). Different isomers can eliminate CO (M8 via TS8-P4) or CO_2 (M3 via TS3-P3 and M5 via TS5-P5) or break into DBB and an open isomer of MA (M6 via TS6-P2). Note that all these paths from M1 have high activation barriers (TS1-P1: (a) 71.9 at the M06-2X/6-31G* level of theory and (b) 64.7 kcal/mol at the B3LYP/6-31G* level of theory; TS1-2: (a) 85.9 and (b) 79.0 kcal/mol and TS1-6: (a) 70.1 and (b) 64.7 kcal/mol) and are predicted to be all less favorable than the backward Diels-Alder reaction (TS_DA: (a) 57.7 and (b) 53.25 kcal/mol).

The cationic Diels-Alder product $M1^+$ can eliminate CO_2 forming $P1^+$ via TS1-$P1^+$ or undergo two different isomerizations (figure 3.5(a)): (i) in blue, the migration of bromine and the transition from a 6-membered ring to a 5-membered ring via TS1-2^+ to form $M2^+$ followed by different bromine (TS2-3^+ and TS3-4^+), hydrogen (TS4-5^+ and TS5-6^+) and -OH (TS6-7^+) migrations ultimately leading to $M9^+$ or (ii) in orange, a hydrogen atom migration (TS1-10^+) followed by a migration of bromine that causes the closure of the 6-membered ring into a 5-membered ring via TS10-11^+ and subsequent bromine (TS11-12^+ and TS12-13^+), hydrogen

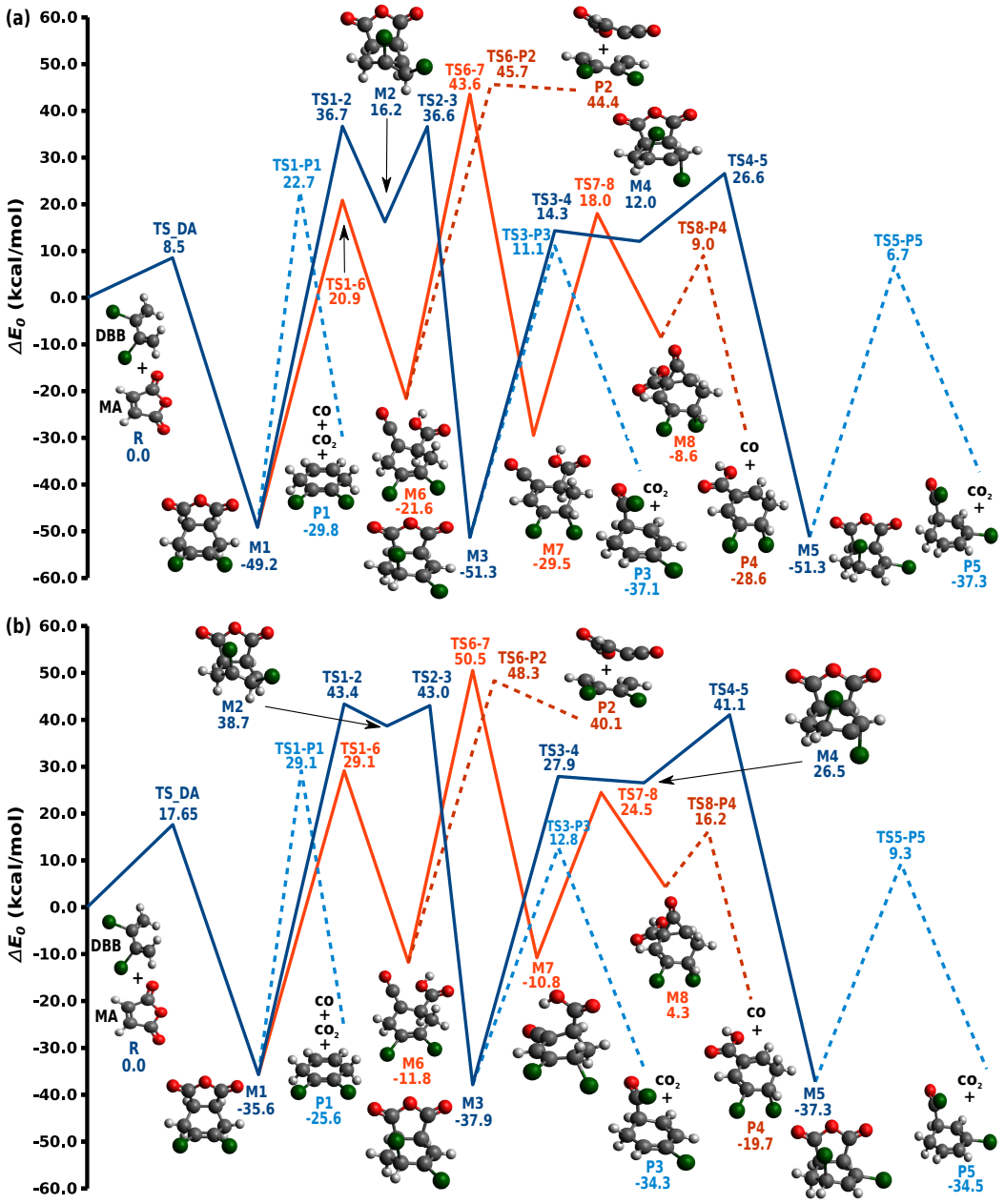


Figure 3.4: Energy profiles at the (a) M06-2X/6-31G* and (b) B3LYP/6-31G* levels of theory for some of the possible isomerizations and fragmentations of the neutral Diels-Alder product. The structures of the different minima are shown. Dashed lines indicate dissociations and solid blue and orange lines two possible isomerization pathways. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

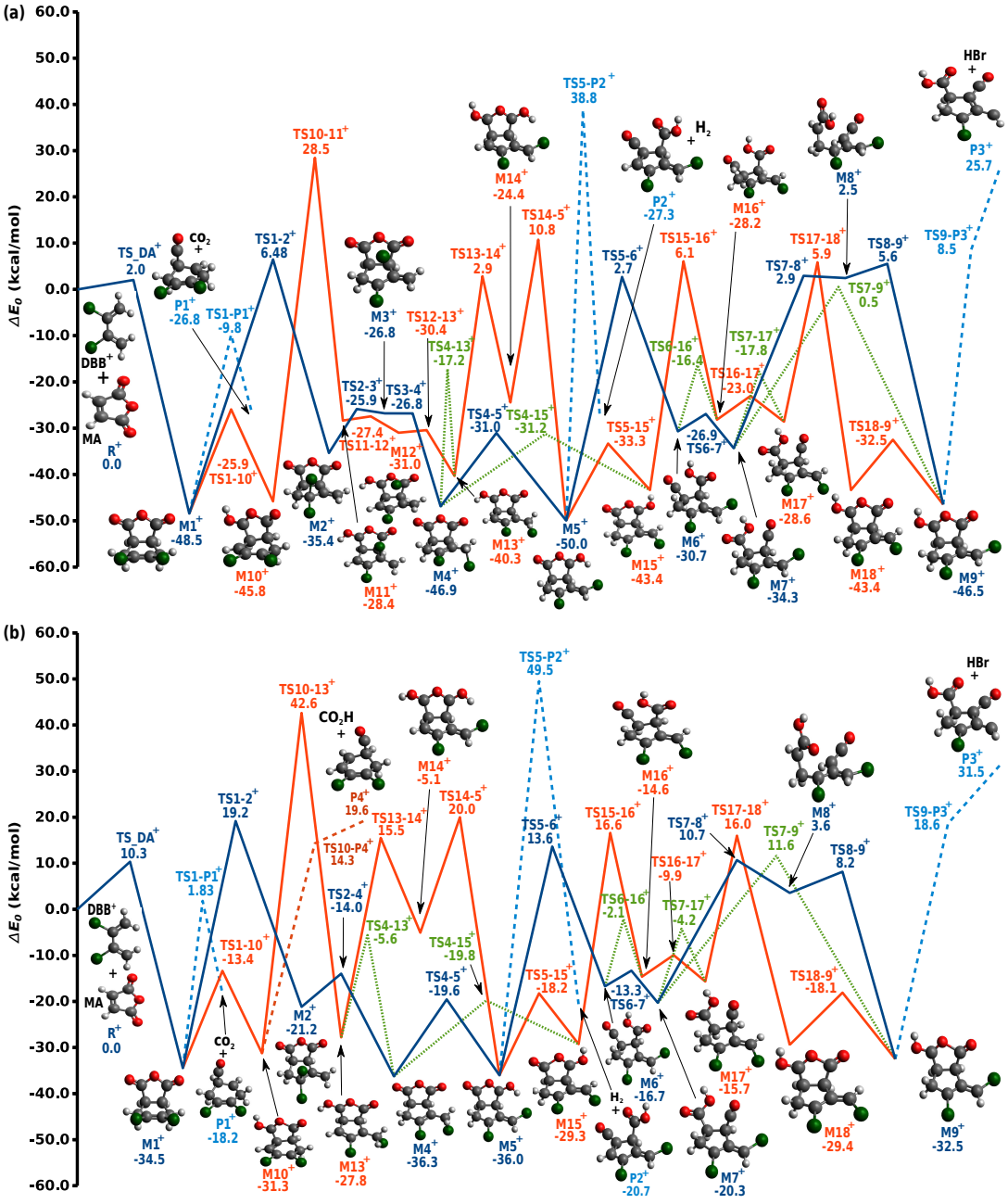


Figure 3.5: Energy profiles at the (a) M06-2X/6-31G* and (b) B3LYP/6-31G* levels of theory for some of the possible isomerizations and fragmentations of the cationic Diels-Alder product. The structures of the different minima are shown. Dashed lines indicate dissociations, solid blue and orange lines two possible isomerization pathways and green lines connections between the two paths. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

(TS13-14$^+$, TS14-5$^+$, TS15-16$^+$ and TS17-18$^+$) and -OH (TS16-17$^+$) migrations ultimately leading again to M9$^+$. In this case, both paths are interconnected. Note that Bouchoux et al. have already found a closure of a 6-membered ring into a 5-membered ring when they explored the evolution of cationic cyclohexene [77]. The system can eliminate H_2 (TS5$^+$-P2$^+$) and HBr (TS9$^+$-P3$^+$) at different points along the isomerization paths (M5$^+$, M9$^+$). In this case, the processes leading to the formation of M10$^+$ (E_a = (a) 22.6, (b) 21.1 kcal/mol) and the dissociation into CO_2 + P1$^+$ (E_a = (a) 38.7, (b) 36.3 kcal/mol) are predicted to be more favorable than the backward Diels-Alder reaction (E_a = (a) 50.5, (b) 44.8 kcal/mol).

The energy profiles of both the neutral and the cationic systems at the B3LYP level (figures 3.4(b) and 3.5(b)) qualitatively agree with the M06-2X picture although for the cationic case some of the shallow intermediates and TSs are not found at the B3LYP level as is the case for M3$^+$, M11$^+$, M12$^+$ amd the TSs that connect them with the rest of the surface.

Comparing the neutral and cationic profiles, it can be seen that the cationic activation barriers are much lower than the neutral ones so the kinetics should be faster in the cationic system. Contrary to the neutral, in the cationic case there are predicted to be two transition states with lower activation barriers than the backward Diels-Alder reaction leading to alternative decay routes of the Diels-Alder products in a gas-phase experiment.

3.3.3 Kinetic calculations

The population of selected species along the previously discussed isomerization paths at different excess energies as a function of the logarithm of time are shown in figures 3.6 and 3.8 for the neutral and cationic reactions, respectively. At t = 0 s, the system is initialized in the product of the Diels-Alder reaction i.e. M1 for the neutral system and M1$^+$ for the cationic one. The populations at each time are calculated by solving equation (3.1). Dissociation processes are treated as irreversible.

The RRKM calculations for the neutral system at 50 kcal/mol of excess energy above the energy of the reactants (i.e. DBB + MA) is shown in figure 3.6(a). We have chosen this particular energy because it is achievable under typical ex-

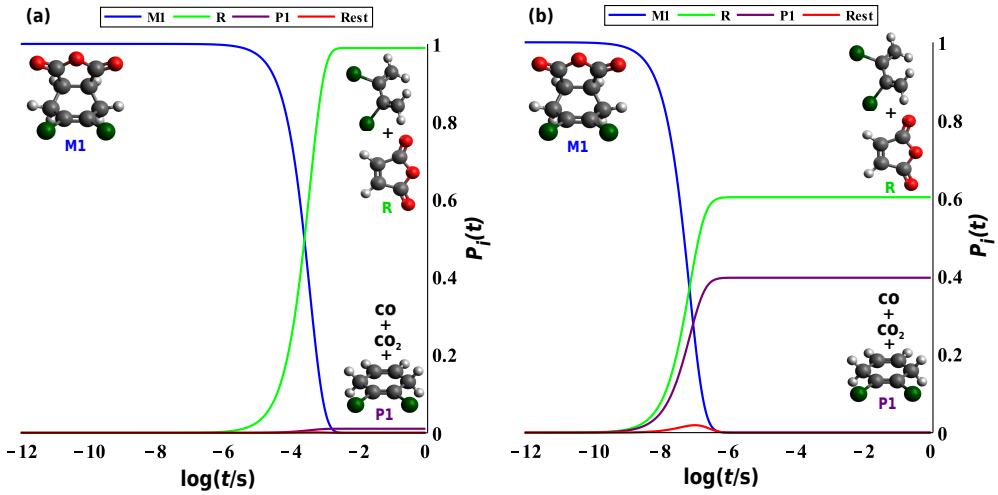


Figure 3.6: RRKM calculations for the neutral system with energies and frequencies calculated at the M06-2X/6-31G* level of theory at an energy of (a) 50 kcal/mol and (b) 100 kcal/mol above the energy of the Diels-Alder reactants assuming a microcanonical ensemble in a collisionless regime. Population of the most important species ($P_i(t)$) versus logarithm of time. Rest corresponds to the sum of the populations of the species that are not explicitly shown. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

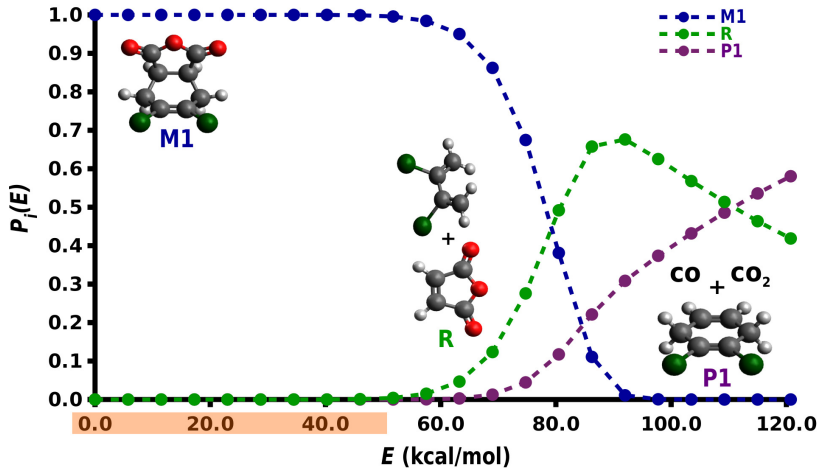


Figure 3.7: RRKM calculations for the neutral system with energies and frequencies calculated at the M06-2X/6-31G* level of theory at $t = 1 \cdot 10^{-6}$ s. Colored dots represent the population of the most important species ($P_i(E)$) versus excess energy in kcal/mol above the energy of the reactants. The dashed lines help to guide the eye. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms. The brown rectangle indicates typical achievable energies in molecular-beam experiments.

perimental conditions. It can be seen that with increasing time the reactants are re-populated which means that the backward reaction dominates the dynamics of M1 at this energy. There is also a small contribution of the dissociation into P1 + CO + CO_2. The isomerization paths are found to be insignificant. This situation changes at double excess energy (100 kcal/mol, figure 3.6(b)), where the kinetics are predicted to be four orders of magnitude faster than in figure 3.6(a). There is now a considerable probability that the system dissociates into P1 + CO + CO_2. Again, the backward reaction plays an important role and the isomerization paths barely participate.

Figure 3.7 shows the distribution of the population of the most important species at $t = 1{\cdot}10^{-6}$ s as a function of the excess energy. This time corresponds to typical experimental timescales for the detection of reaction products in molecular-beam experiments. It can be seen that the product of the Diels-Alder reaction M1 is stable on the microsecond timescale at experimentally achievable energies (indicated in brown in figure 3.7) and only at high excess energies dissociation products are obtained on this timescale. Even at the highest energies studied, no noticeable participation of the isomerization pathways was found.

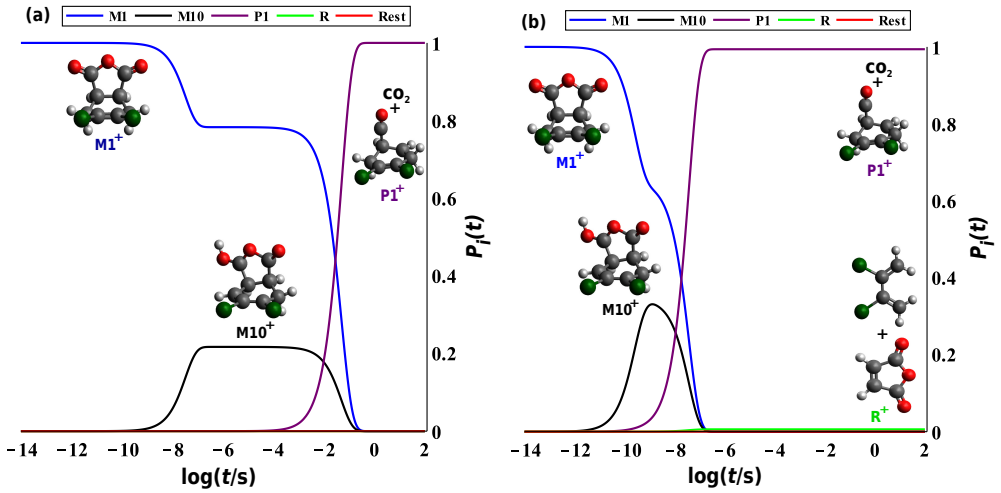


Figure 3.8: RRKM calculations for the cationic system with energies and frequencies calculated at M06-2X/6-31G* level of theory at an energy (a) equal to 0.0 kcal/mol and (b) equal to 45.34 kcal/mol above the energy of the Diels-Alder reactants assuming a microcanonical ensemble in a collisionless regime. Population of the most important species ($P_i(t)$) versus logarithm of time. Rest corresponds to the sum of the populations of the species that are not explicitly shown. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

The evolution of the cationic reaction at the energy of the reactants (i.e. DBB^+ + MA) is shown in figure 3.8(a). At this available energy, the open channels are:

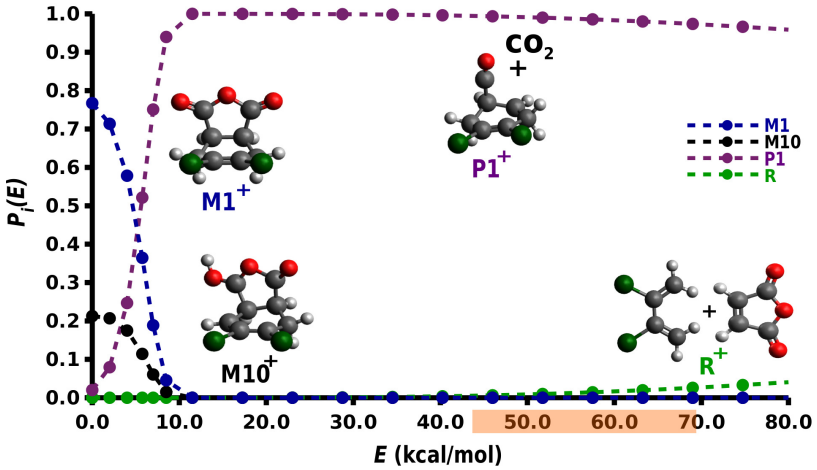


Figure 3.9: RRKM calculations for the cationic system with energies and frequencies calculated at M06-2X/6-31G* level of theory at $t = 1 \cdot 10^{-3}$ s. Colored dots represent the population of the most important species ($P_i(E)$) versus excess energy in kcal/mol above the energy of the reactants and the dashed lines are for eye guidance. Green spheres represent bromine atoms, red spheres oxygen atoms, gray spheres carbon atoms and white spheres hydrogen atoms. The brown rectangle indicates typical experimentally achievable energies assuming that the charge is initially located on MA.

dissociation into $P1^+ + CO_2$, isomerization into $M10^+$ and the backward reaction from $M10^+$ to $M1^+$. We first see how part of the population of $M1^+$ isomerizes into $M10^+$ until an equilibration between the populations of the two species is reached. Due to the fact that the most favorable process to happen from $M1^+$ is the isomerization into $M10^+$ while the other possible transitions from $M1^+$ and $M10^+$ have much higher activation energies, the system is trapped in these two minima until it overcomes the barrier for dissociating into $P1^+ + CO_2$. Therefore, with this available energy and on typical timescales of experiments, both isomerization and dissociation into $P1^+ + CO_2$ of the cationic Diels-Alder product are predicted. The RRKM results of $M1^+$ at an excess energy of 45.34 kcal/mol are displayed in figure 3.8(b). This energy corresponds to the energy of the experimentally achievable configuration DBB + MA^+, i.e. the configuration in which the charge has initially been exchanged between the reactants relative to the energy of the reactants DBB^+ + MA, at the M06-2X/6-31G* level. The reaction seems to be much faster than without excess energy. The time during which the system is trapped between $M1^+$ and $M10^+$ is now much shorter. The main channel is again the dissociation into $P1^+ + CO_2$, but there is now a small contribution of the backward Diels-Alder reaction to the kinetics. At this energy, dissociation into $P1^+ + CO_2$ is observed on the nanosecond timescale which should be detectable in experiments. Note, however, that the asymptote DBB + MA^+ corresponds to

an excited state of the system so excited-state dynamics may play a significant role in this case.

To further explore the reactivity of $M1^+$, figure 3.9 shows how the populations of the different species in figure 3.5 change as a function of excess energy at $t = 1 \cdot 10^{-3}$ s which corresponds to a typical experimental timescale for reaction experiments with ions in traps. It can be seen that at low energies, the isomerization into $M10^+$ is the dominant reaction of $M1^+$. At increasing energy, the main decay channel is the dissociation into $P1^+ + CO_2$. At the highest energies studied, the backward Diels-Alder reaction starts to participate.

3.4 Conclusion

The neutral and cationic Diels-Alder reactions between DBB and MA have been studied at the DFT level of theory. The neutral reaction is predicted to be concerted and symmetric while the cationic reaction is asymmetric and can either be concerted or stepwise. The s-*trans* conformer of DBB can contribute to the cationic, stepwise reaction. The time evolution of possible products of these reactions until dissociation under typical experimental conditions has also been explored. The neutral system has higher activation barriers than the cationic system. Kinetic studies based on RRKM theory indicate that the kinetics of the cationic system is much faster than that of the neutral system due to lower activation energies. The neutral product is predicted to most likely return to the reactants on a timescale of hundreds of microseconds at an excess energy of 50 kcal/mol, which corresponds to an achievable excess energy in experiments. On the other hand, the cationic product is expected to eliminate CO_2 on a nanosecond timescale at an excess energy of 45.34 kcal/mol, which approximately corresponds to the conditions in which MA is asymptotically the cationic species.

As a next step, the dynamics of this reaction are explored in chapter 4 for the neutral system and in chapter 5 for the cationic variant by means of reactive molecular dynamics. In this way questions about the synchronicity of the process and possible ways of promoting the reaction can be studied.

To our knowledge, this specific system had not been studied until now. Gaining insight into the differences between the cationic and neutral systems is crucial to understand the enhanced selectivity and rates of the cationic Diels-Alder re-

actions compared to their neutral counterparts. The present results may serve to design single-collision experiments that will probe these reactions and aid in their interpretation.

Chapter 4

Reactive Molecular Dynamics of DBB + MA*

4.1 Introduction

Here, the gas-phase neutral Diels-Alder reaction between DBB and MA (see scheme 3.1) is studied by means of reactive molecular dynamics simulations. In chapter 3, it was shown that the Diels-Alder product will only fragment on the nanoseconds timescale at the highest energies studied and it will most certainly fragment back to the reactants. For this reason, only the Diels-Alder part of the potential energy surface (see figure 3.1) will be taken into account for the dynamic study. Besides, since atomistic simulations typically cover picoseconds timescales, we will only study the forward reaction, i.e. the formation of products and not the fragmentation of the product back to the reactants. The trajectories start with the two reactant molecules approaching each other in order to simulate a collision experiment. This allows us to address, in an unbiased manner, questions such as whether the reaction is synchronous, whether the mechanism is complex-mediated and how this reaction could be promoted.

The parametrization of the models used for the simulations is explained in sections 4.2.2 and 4.2.3 and their quality shown in section 4.3.1. The minimum

*This chapter is based on the paper *Reactive Atomistic Simulations of Diels-Alder reactions: the Importance of Molecular Rotations.* U. Rivero, O. T. Unke, M. Meuwly, S. Willitsch, Journal of Chemical Physics 151 (2019) 104301.

dynamic path is analyzed in section 4.3.2, the van der Waals complex formation in the entrance channel is explored in section 4.3.3 and reactive trajectories are obtained in section 4.3.4.

4.2 Methods

4.2.1 Molecular Dynamics Simulations

Atomistic simulations were carried out either with the CHARMM program [129] when using MS-ARMD [99] based on an initial parametrization from SwissParam [130] or with the Atomic Simulation Environment (ASE) [131] when using the PhysNet neural network architecture designed for predicting energies, forces and dipole moments of chemical systems [103].

All bonds involving hydrogen atoms were flexible and the time step used in the simulations was $\Delta t = 0.1$ fs to ensure conservation of total energy. The velocity Verlet algorithm was used for the propagation of the equations of motion [114].

The initial SwissParam parametrization was modified in order to construct a multisurface adiabatic reactive molecular dynamics (MS-ARMD) [99] force field for the present Diels-Alder reaction. Ensembles of structures for the parametrization of the MS-ARMD model were generated with CHARMM as follows: the optimization of the system with the adopted Newton-Raphson method was followed by 50 ps of heating dynamics, 50 ps of equilibration at 500 K, 60 ps of cooling down to 300 K and free *NVE* (microcanonical ensemble) dynamics. The temperature was only raised up to 400 K for parts of the parametrization of the reactant van der Waals complex to avoid dissociation.

For the PhysNet parametrization, initial ensembles for the different fragments [102] were generated using the PM7 level of theory implemented in MOPAC2016 [93, 132] and subsequently augmented using adaptive sampling [117, 133].

In order to generate the initial conditions for the collision simulations, ensembles of the individual molecules (MA and DBB) at different vibrational temperatures were generated using CHARMM as described above. Heating and equilibration temperatures were modified depending on the desired final vibrational temper-

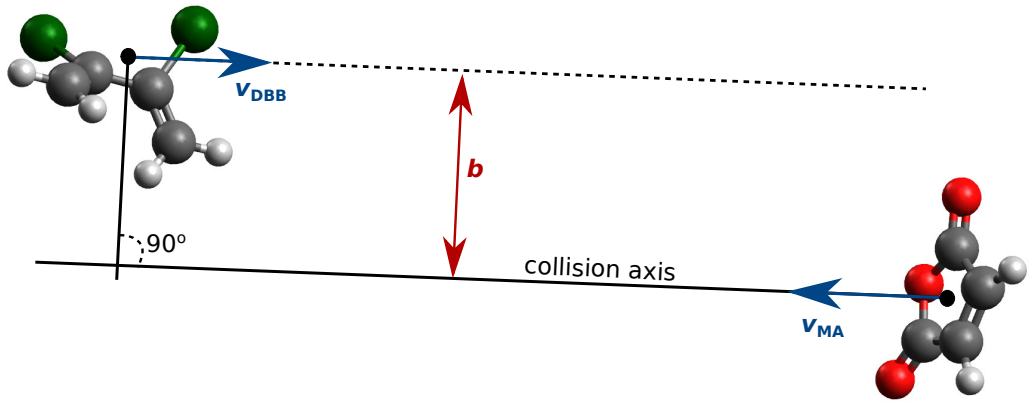


Figure 4.1: Schematics of the initial conditions of a trajectory. The reactant molecules are set at an initial distance of 20 Å along the collision axis and the impact parameter (b) in red is set by displacing DBB along a perpendicular axis. The blue arrows represent the initial velocities of the center of mass of DBB (v_{DBB}) and MA (v_{MA}).

ature (T_{vib}). The reactants were placed at an initial distance of 20 Å (taking the center of mass of each molecule as reference) with a random relative orientation. In order to tune the collision energy (E_{coll}), the atomic velocities along the collision axis were modified. The impact parameter (b) was uniformly sampled by displacing one of the molecules along a perpendicular axis (see figure 4.1). Rotational temperature (T_{rot}) was added following calculation of the moment of inertia tensor and assuming equipartition among the three rotational degrees of freedom [134]. Excitation of specific vibrational normal modes was achieved by projecting the initial velocities onto the space of normal modes and modifying the kinetic energy of the desired normal mode.

The trajectories were considered reactive when the C-C distance between the carbon atoms involved in the two new bonds was smaller than 1.6 Å. The reactive cross section σ was calculated as stated in equation (2.78) in chapter 2. The reaction rate k can be calculated from

$$k = \sigma \cdot v_{\mathrm{rel}}, \tag{4.1}$$

where v_{rel} is the relative center of mass velocity of the colliding molecules [116].

The total kinetic energy of the minimum-dynamic-path trajectories was decomposed either by projecting it onto the degrees of freedom of the system, where the reference degrees of freedom were calculated as the eigenvectors of the Hessian matrix of the isolated, reactant molecules with geometries corresponding to

the last point of each trajectory; or into the translational energy of the center of mass of the reactant molecules (E_{trans}) the rotational energy corresponding to their angular momentum (E_{rot}) and vibrational energy (E_{vib}) as stated in equations (4.2) to (4.7):

$$E_{\text{trans},j} = \frac{|\sum_{i\in j} \vec{p}_i|^2}{2 \cdot M_j}, \tag{4.2}$$

where $\vec{p}_i$ is the momentum of atom i belonging to molecule j (j = MA, DBB) and M_j is the total mass of molecule j;

$$E_{\text{rot},j} = \frac{1}{2}|\mathbf{I}_j \vec{\omega}_j^2|, \tag{4.3}$$

where $\vec{\omega}_j$ is the angular velocity of molecule j and $\mathbf{I}_j$ is the inertia tensor of molecule j:

$$\vec{\omega}_j = \mathbf{I}_j^{-1} \vec{L}_j. \tag{4.4}$$

$\vec{L}_j$ is the angular momentum of molecule j:

$$\vec{L}_j = \sum_{i\in j} \vec{r}_i\,' \times \vec{p}_i\,', \tag{4.5}$$

where atomic momenta ($\vec{p_i}\,'$) and atomic coordinates ($\vec{r_i}\,'$) in the center of mass frame are calculated as:

$$\vec{x}_i\,' = \vec{x}_i - \vec{x}_{\text{CoM},j}; \qquad x = p, r \tag{4.6}$$

and the subscript CoM, j refers to the center of mass of molecule j.

$$E_{\text{vib},j} = E_{\text{tot},j} - E_{\text{rot},j} - E_{\text{trans},j} \tag{4.7}$$

where $E_{\text{tot},j}$ is the total kinetic energy of molecule j along the trajectory.

4.2.2 Parametrization of MS-ARMD

The force fields for the reactant and product states were parametrized to reproduce DFT reference energies at the M06-2X/6-31G* level of theory which was found to yield the best accuracy at the DFT level for this type of reaction by Linder and Brinck [63, 98] and following our previous work [135]. All single point calculations were performed using Gaussian09 [121]. Starting from parameters retrieved from SwissParam [130], first ensembles for reactants and products were generated as described in section 4.2.1. Reference energies were calculated and the individual force fields were fitted using a downhill simplex algorithm [117]. The standard CHARMM harmonic potentials for the description of C-C, C=C, C-O and C-Br bonds were replaced by Morse potentials (C-H and C=O bonds were kept as harmonic). Furthermore, in the reactant force field, the Lennard-Jones potentials between the four carbon atoms involved in the Diels-Alder reaction as well as between Br/O atoms were replaced by a generalized Lennard-Jones potential [99]. The remaining terms were parametrized as in the standard CHARMM force field.

With the initial parametrization of the individual force fields, another ensemble of structures was generated for reactants and products with the new parameters and added to the training structures. The reference energies were calculated and the parametrization was further refined. This iterative procedure continued until the root-mean-square deviation (RMSD) of the newly generated ensemble was approximately the same as that of the previous iteration of the parametrization. For the reactant force field, 401 and 2064 structures where required for the parametrization of MA and DBB, respectively. The non-bonded terms of the van der Waals complex were parametrized with 2783 additional structures. For the parametrization of the product force field, 2589 structures were required.

The crossing region between the two force fields was smoothed following the internal reaction coordinate (IRC) by combining the final force fields for reactants and products with GAussian times POlynomial functions (GAPOs) [99]. A genetic algorithm was used for the fitting of the GAPOs. The global PES is given

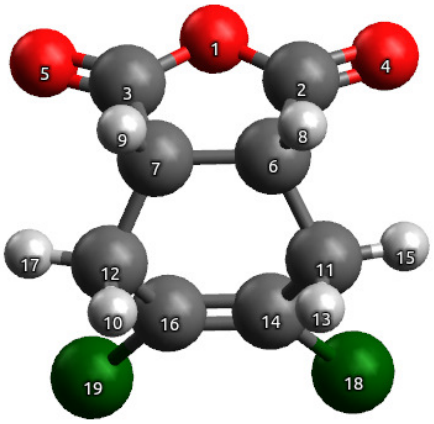


Figure 4.2: Numbering of the atoms for the force fields for one of the two product force fields. The other force field has the dibromobutadiene rotated such that the C-C bonds are between atoms 6-12 and 7-11. Carbons are in gray, hydrogens in white, oxygens in red and bromines in green.

by equation (4.8) where $V_i(x)$ are the individual force fields, $w_i(x)$ their weights and $\Delta V^{ij}_{\mathrm{GAPO},k}(x)$ the GAPOs [99].

$$V_{\mathrm{MS-ARMD}} = \sum_{i=1}^{n} w_i(x) V_i(x) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} [w_i(x) + w_j(x)] \sum_{k=1}^{n_{ij}} \Delta V^{ij}_{\mathrm{GAPO},k}(x) \qquad (4.8)$$

The product of the Diels-Alder reaction has two possible connectivities (see figure 4.2). In order to make the parametrization of the product permutation invariant, two force fields that describe these two possible connectivities were used.

4.2.3 Parametrization of PhysNet

Reference data for training PhysNet (energies, forces and dipole moments) were calculated at the DFT M06-2X/6-31G* level of theory using Gaussian09 [121]. All possible "amon" fragments [102] for DBB, MA, and their reaction product were generated (378 in total) and different geometries for all fragments were sampled by running Langevin dynamics at 1000 K at the PM7 level of theory. After training PhysNet models on this initial dataset, additional structures were generated by adaptive sampling [133, 136]: an ensemble of 4 PhysNet models was used to run Langevin dynamics at 1000 K and new *ab initio* data was calculated

for geometries for which the energy predictions between the different models differed by more than 0.5 kcal/mol. For further details on the adaptive sampling method, see Ref. [103]. The dataset was iteratively augmented in this fashion until no significant deviations between the predictions of individual PhysNet models could be observed (the final dataset contains 224483 structures [137] which is approximately 50 times larger than it was required for the MS-ARMD model). PhysNet was then trained on 200000 structures of the final dataset (with 20000 additional structures used for validation) by minimizing the mean absolute error between neural network predictions and the reference data using the procedure given in Ref. [103] (all hyperparameters of the neural network architecture and the training procedure were set to the values recommended in Ref. [103]). The global PES is given by

$$V_{\text{PhysNet}} = \sum_{i=1}^{N} E_i + k_e \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{q_i q_j}{r_{ij}} \tag{4.9}$$

where N is the total number of atoms, k_e is the Coulomb constant, r_{ij} is the distance between atoms i and j, and E_i and q_i are atomic energy contributions and partial charges (corrected to guarantee charge conservation [103]) predicted by PhysNet. Here, the Coulomb potential is damped at short distances to avoid numerical problems (see Ref. [103]). The PhysNet architecture guarantees that Eq. 4.9 is invariant with respect to translations, rotations and permutation of atoms sharing the same element type [103].

4.3 Results and Discussion

4.3.1 Parametrization of the Reactive Force Fields

Figure 4.3 shows stationary points on the PES for the Diels-Alder reaction between DBB and MA at the M06-2X/6-31G* level of theory. Note that DBB is in a *gauche* conformation (the C=C−C=C dihedral angle has a value of 50°) since the s-*cis* geometry is not a minimum on the PES. Due to the fact that both reactant molecules are symmetric, there are only two possible paths for the Diels-Alder

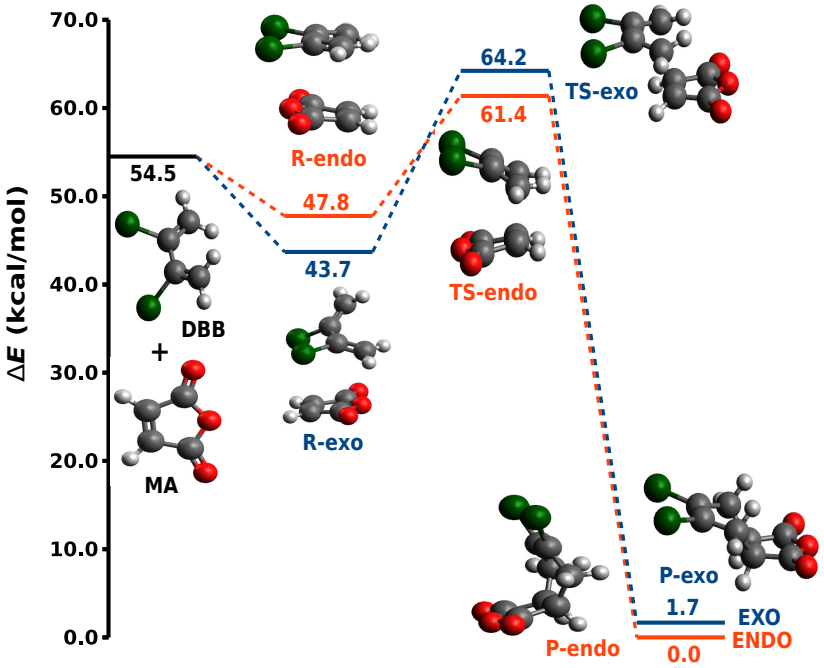


Figure 4.3: Potential energy surface for the two possible Diels-Alder reaction paths (exo and endo in blue and orange, respectively) between dibromobutadiene (DBB) and maleic anhydride (MA) at the M06-2X/6-31G* level of theory. The relative energies in kcal/mol with respect to the endo product (P-endo) as well as the structures of minima and transition states are shown.

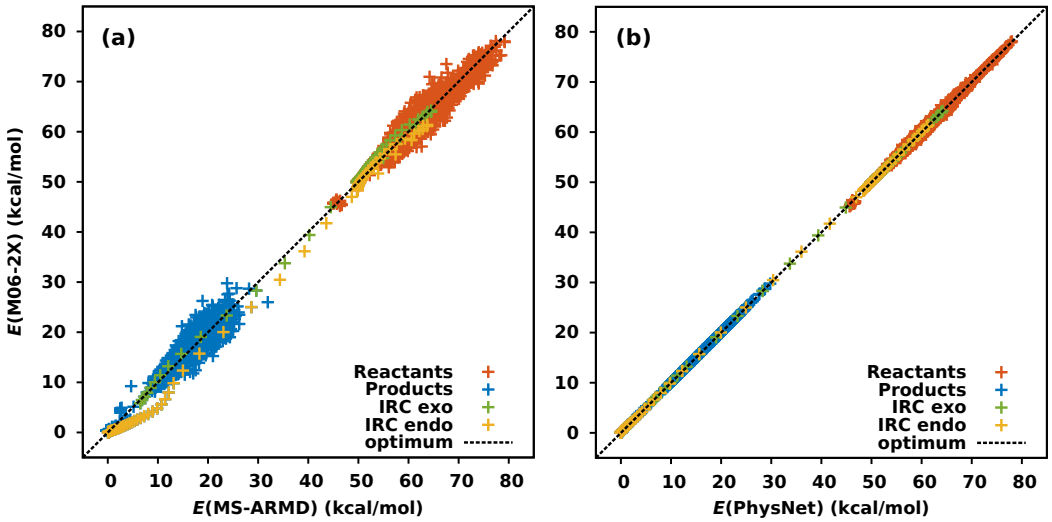


Figure 4.4: Energy correlation of 5512 M06-2X/6-31G* reference structures and (a) the MS-ARMD model with a total RMSD of 1.47 kcal/mol and $R^2 = 0.9961$ or (b) the PhysNet model with a total RMSD of 0.25 kcal/mol and $R^2 = 0.9999$.

reaction. These are referred to as "endo" and "exo" depending on the relative orientation of the reactants. The endo product (P-endo) is taken as the zero of the energy scale in figure 4.3. Normally, the intermolecular interactions favor the endo path which M06-2X correctly describes [127, 135, 138]. The dissociation of the van der Waals complexes (R-endo and R-exo) towards the reactants is more favorable than the reaction over the barrier towards the Diels-Alder products (see figure 4.3). Judging from the PES and the geometries of the TSs, the reaction should be concerted and symmetric.

As discussed in the introduction, in order to answer questions such as whether the reaction is synchronous or complex mediated, the dynamics of the system must be studied. Therefore, we constructed a computationally efficient PES that allows us to run a statistically significant number of trajectories, such that diverse initial conditions can be sampled. The quality of the parametrized MS-ARMD model is shown in figure 4.4(a). Points generated during the parametrization of the product force field (2589 structures) and the parametrization of the non-bonded interactions of the reactant force field (2783) are shown, as well as the IRCs for the endo (81) and the exo (59) paths. The exo IRC was used for the parametrization of the GAPOs and thus it is better described by the model than the endo path. The energy of the endo IRC is slightly overestimated by the MS-ARMD model. The total RMSD is 1.47 kcal/mol over a range of 80 kcal/mol which we asses to be sufficient for an adequate characterization of the dynamics of the system. It is important to note that the s-*trans* conformer has not been included and is not stable in this model. Thus, the reactants will not sample the entire conformational space leading to a possible overestimation of the reaction rates since the s-*trans* conformer does not contribute to the reaction i.e. the number of unreactive trajectories increases when including the s-*trans* conformer which leads to a reduction of the reaction flux and hence the rate. For the validation of the MS-ARMD model the reactant, transition state and product structures were minimized and compared to the DFT geometries in figure 4.5 where the main difference between the minimized structures is the C=C−C=C dihedral angle of DBB in the van der Waals complex (figure 4.5(a)). Describing a weakly bound complex with point charges and isotropic van der Waals interactions is an approximation. The structure of the van der Waals complex can be further improved by using multipolar interactions for the electrostatics and jointly fitting van der Waals and dihedral parameters. However, this was not further pursued

because the correlation between reference and fitted energies is sufficiently good and structural details are less relevant for this work.

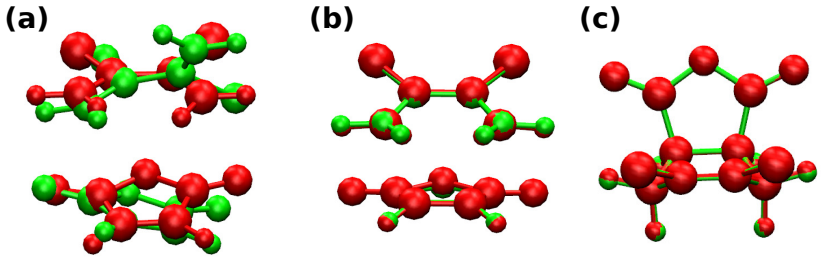


Figure 4.5: Comparison of (a) reactant, (b) transition state and (c) product minimized structures of the endo path at DFT (red) and MS-ARMD (green) levels of theory.

The harmonic frequencies have also been compared between DFT and the MS-ARMD model in figure 4.6. The parametrized model is given in tables A.1-A.7 of appendix A.

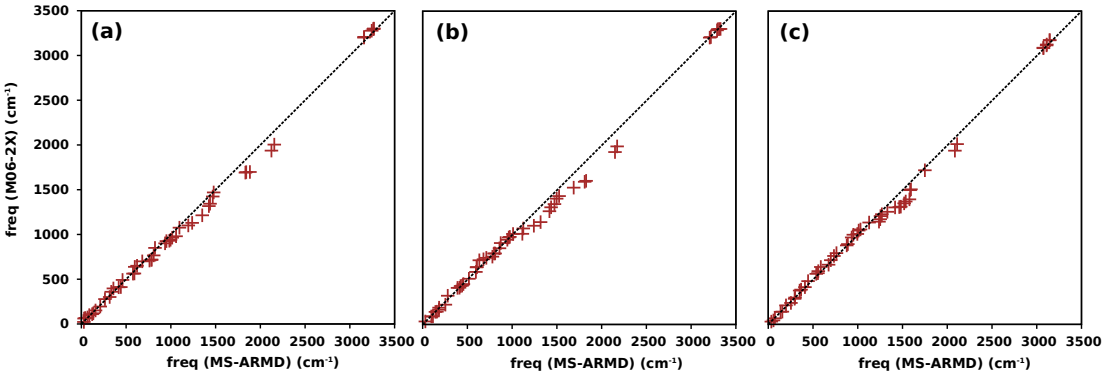


Figure 4.6: Comparison of (a) reactant, (b) transition state and (c) product harmonic frequencies for the minimized structures in figure 4.5 at DFT and MS-ARMD levels of theory with RMSDs of 69.4, 85.0 and 68.5 cm^{-1} respectively.

For validation and direct comparison, a PhysNet model was parametrized for the same system. Figure 4.4(b) reports on the quality of this PES by showing the correlation between the reference energies and the PhysNet predictions on the training structures of the MS-ARMD model with a total RMSD of 0.25 kcal/mol. The PhysNet model is significantly more accurate than MS-ARMD, but it has a computational cost 200 times higher.

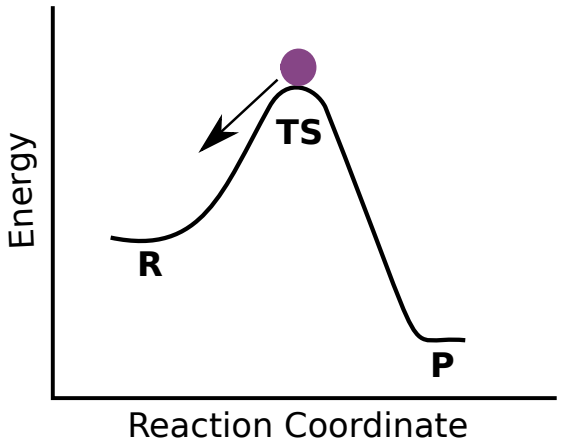


Figure 4.7: Schematics of a minimum-dynamic-path trajectory towards reactants. The purple dot represents the initial conditions of the simulation. The system starts at the transition state (TS) with zero atomic velocities and gains kinetic energy as it roles down to reactants (R).

4.3.2 Minimum Dynamic Path

A trajectory starting at the TS geometry without kinetic energy (the total energy of the trajectory is equal to the potential energy of the TS) follows the minimum dynamic path (MDP) [139]. The projection of the total kinetic energy along the MDP towards the reactants (see figure 4.7) onto the degrees of freedom of DBB and MA is shown in figure 4.8(a) as sums of translational, rotational and vibrational energies. At $t = 0$ fs, the system is at the TS and at $t = 100$ fs it has arrived at the reactant state. At the beginning of the trajectory, the system contains no kinetic energy and moves only slowly, until at around $t = 50$ fs more potential energy is converted into kinetic energy. By projecting the total kinetic energy onto the different degrees of freedom of DBB and MA, the active degrees of freedom can be identified. Those degrees of freedom to which most kinetic energy is imparted will be important for driving the system towards the transition state. Figure 4.8(a) shows that rotations contain the largest amount of kinetic energy for both DBB and MA. The rotational energy of DBB and MA together accounts for 63% of the total kinetic energy while translational energy accounts for 18%. Certain vibrations are also activated (see figure 4.9(a) for the individual contributions and figure 4.10 for the active rotations). An identical result has been obtained from the direct decomposition of the total kinetic energy (figure 4.11). Using the sudden vector projection method [140] we also arrive at the conclusion that rotations are the most important degrees of freedom. This is surprising since

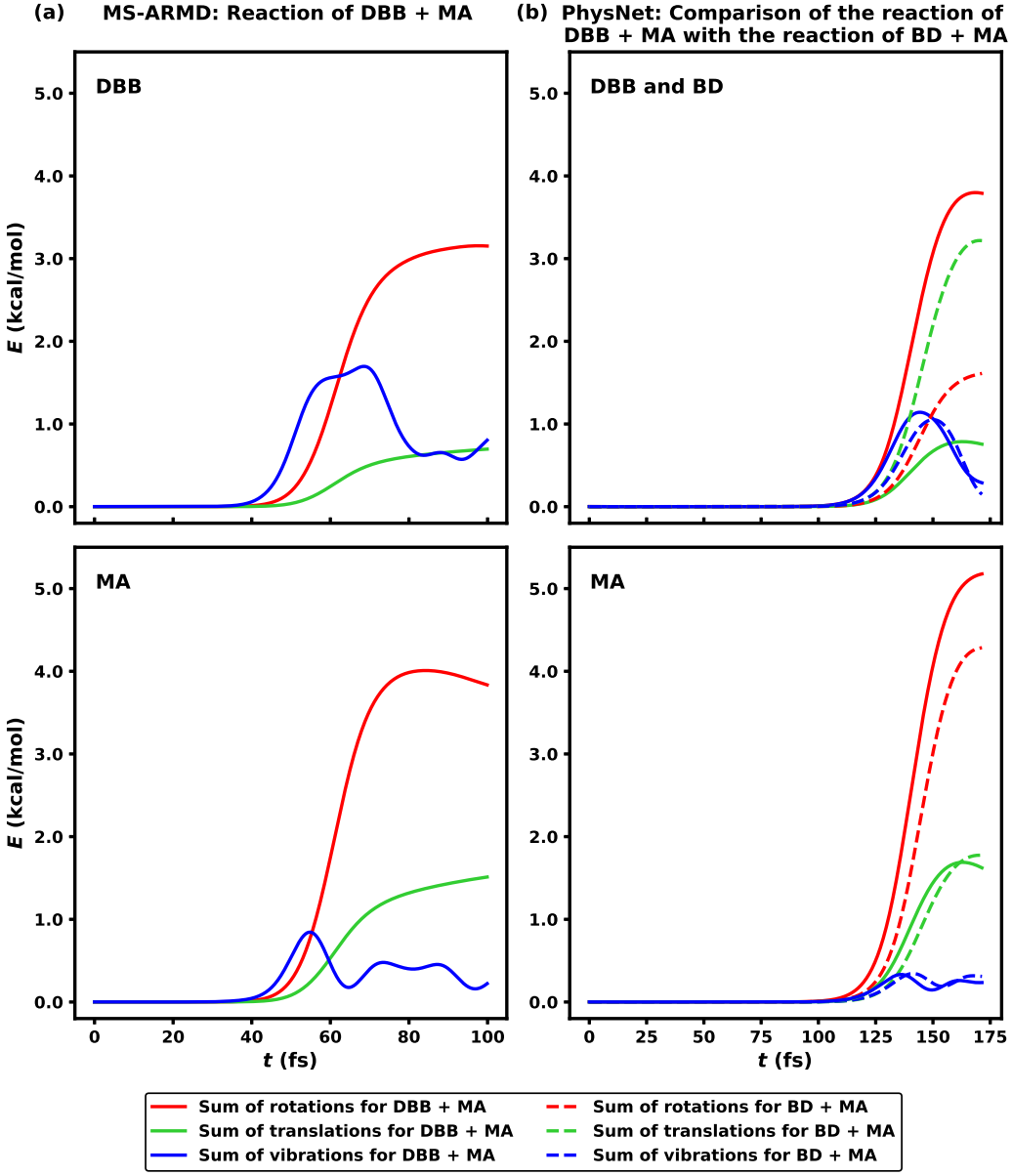


Figure 4.8: Solid lines: projection of the total kinetic energy (E) onto the degrees of freedom (rotations, translations and vibrations) of dibromobutadiene (DBB) and maleic anhydride (MA) for the reaction of DBB + MA along the minimum dynamic path calculated with (a) the MS-ARMD model and (b) the PhysNet model. Dashed lines in panel (b): projection of the total kinetic energy onto the degrees of freedom of butadiene (BD) and MA for the reaction of BD + MA along the minimum dynamic path calculated with the PhysNet model. The trajectories start at the endo transition state and end at the reactants.

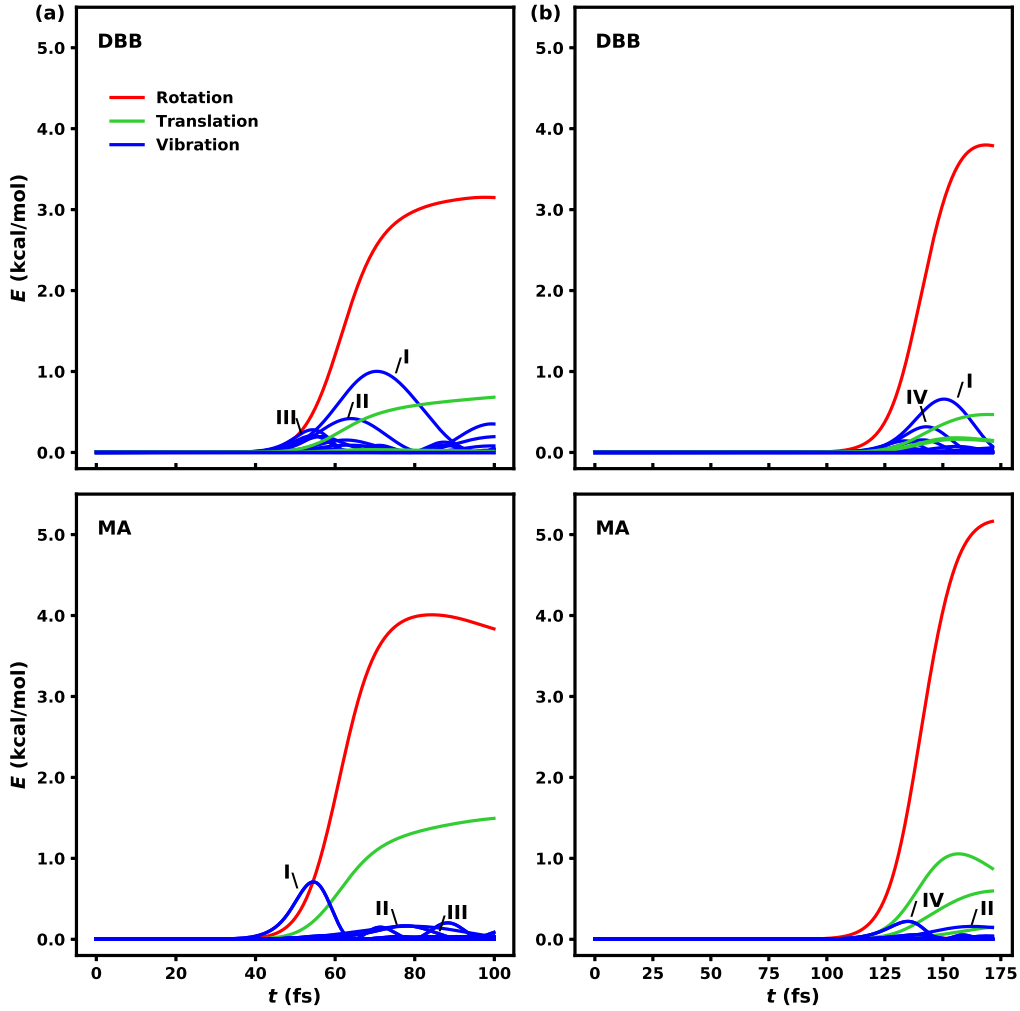


Figure 4.9: Projection of the total kinetic energy (E) onto the degrees of freedom (rotations, translations and vibrations) of dibromobutadiene (DBB) and maleic anhydride (MA) for the reaction of DBB + MA along the minimum dynamic path calculated with (a) the MS-ARMD model and (b) the PhysNet model. The trajectory starts at the endo transition state and ends at the reactants. The most important vibrations for DBB are (I) molecular out-of-plane bending, (II, III) C-H out-of-plane bending mixed with skeleton vibrations and (IV) C=C in-plane symmetric bending. For MA they are (I) C-C symmetric stretching, (II) C=O out-of-plane symmetric bending and (III, IV) out-of-plane bendings of the molecule.

one may naively assume that the reaction coordinate is mainly a translation and not a rotation.

In order to further validate this result, the MDP was also calculated using the PhysNet PES. The results are shown in figure 4.8(b) and figure 4.9(b) and qualitatively agree with those of MS-ARMD (68% of the total kinetic energy is imparted into rotations and 20% into translations), although less kinetic energy is acquired

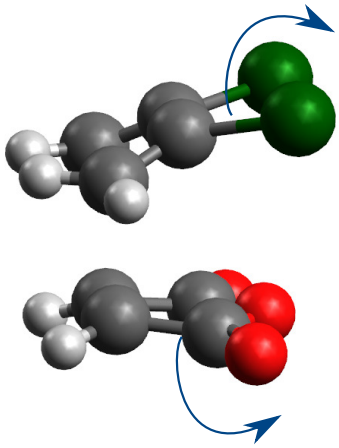


Figure 4.10: Schematic representation of the active rotations of dibromobutadiene and maleic anhydride in the minimum dynamic path starting from the endo transition state.

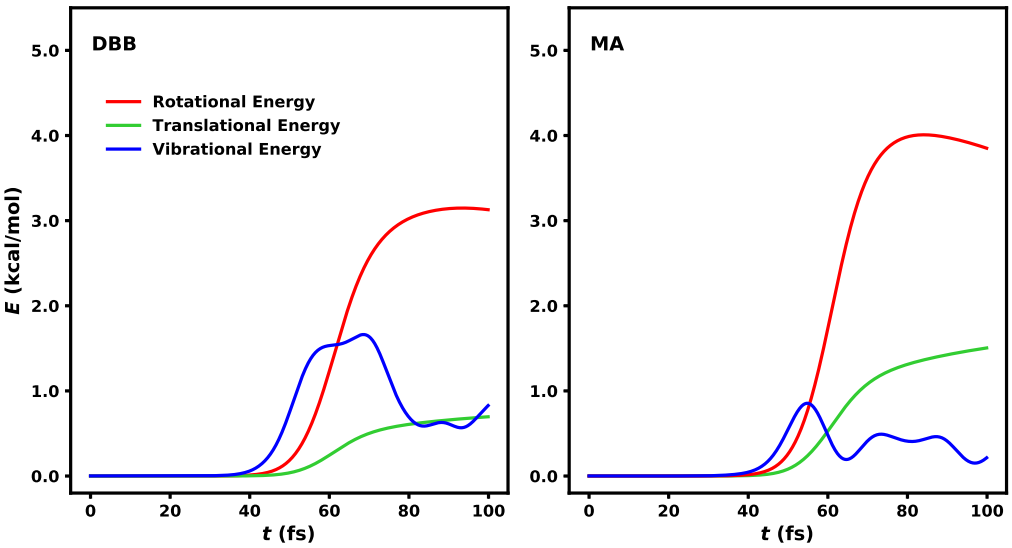


Figure 4.11: Decomposition of the total kinetic energy (E) along the minimum dynamic path calculated with the MS-ARMD model into translational, rotational and vibrational kinetic energy as stated in equations 4.2-4.7. The trajectory starts at the endo transition state and ends at the reactants.

by the vibrations with PhysNet: vibrational energy accounts for 12% of the total kinetic energy while for MS-ARMD it is 19%.

The fact that the MDP between TS and reactant for MS-ARMD extends over 100 fs while for the PhysNet PES it spans for 175 fs is due to differences in the shape of the two PESs near the TS (see figure 4.12) which define the initial accelerations of the particles.

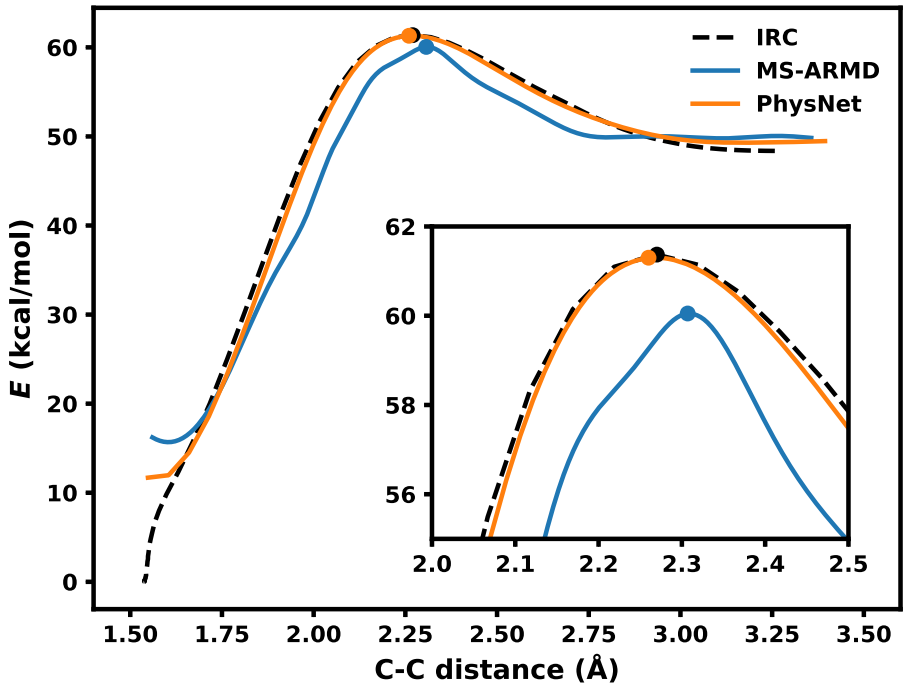


Figure 4.12: Comparison of the shape of the potential energy surface explored by the complete minimum dynamic path performed with MS-ARMD (in blue) and PhysNet (in orange) models as a function of the C-C distance of the bonds formed (the distances of the two bonds formed are identical along the trajectory). The internal reaction coordinate (IRC) calculated at the M06-2X/6-31G* level of theory is also shown as a function of the same coordinate in black. The position of the transition state is marked as a dot. The reactant state is at a C-C distance of 3.35 Å and the product at a distance of 1.6 Å. A magnification of the transition state region is shown in the inset.

In order to investigate whether rotations are important only due to the presence of the heavy bromine atoms in the system (the large mass of bromine atoms potentially enhances the torque on DBB along the MDP), the MDP of the reaction of MA with butadiene (BD) was calculated with the PhysNet model. This is possible since the dataset on which the model was trained contains the necessary information to describe the system with hydrogen atoms instead of bromine

atoms. The results are shown in figure 4.8(b) (individual contributions are shown in figure 4.13). It was found that the important degrees of freedom for MA remain approximately the same for the reaction with DBB and the reaction with BD. When comparing the distribution of kinetic energy of BD with that of DBB, it can be seen that, even though rotations are still active, translations seem to be imparted the most kinetic energy. The rotational energy of BD and MA together accounts in this case for 48% of the total kinetic energy while translational energy accounts for 40%. This indicates that the heavy bromine atoms accentuate the importance of rotational excitation for driving the reaction.

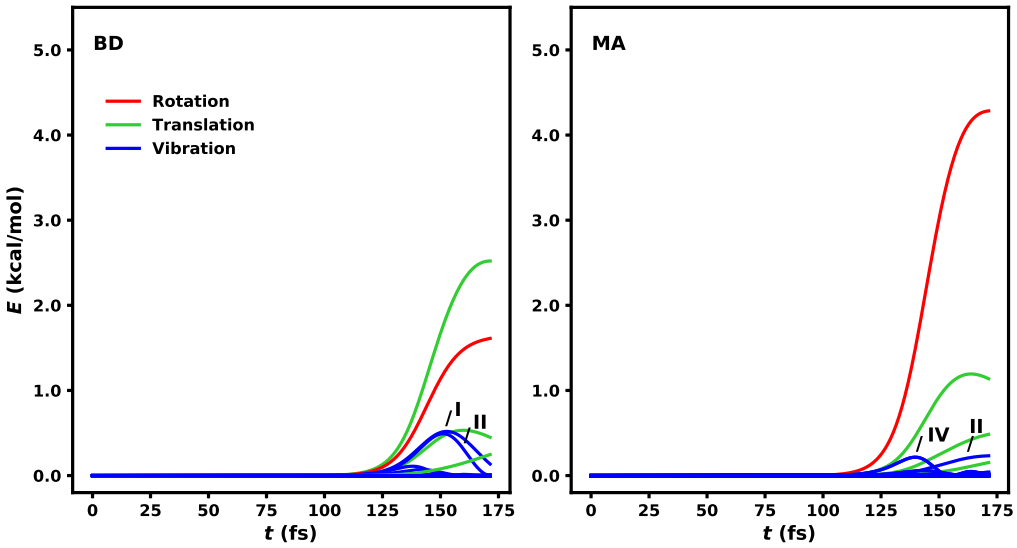


Figure 4.13: Projection of the total kinetic energy (E) along the minimum dynamic path calculated with the PhysNet neural network onto the degrees of freedom of butadiene (BD) and maleic anhydride (MA). The trajectory starts at the endo transition state and ends at the reactant. The most important vibrations for butadiene are (I) C=C in-plane symmetric bending and (II) C-H out-of-plane symmetric bending. For MA they are (II) C=O out-of-plane symmetric bending and (IV) out-of-plane bending of the molecule.

4.3.3 Cross section for the formation of van der Waals complexes in the entrance channel

The formation of van der Waals complexes in the entrance channel was studied in order to determine whether the reaction is direct (without the formation of complexes) or complex-mediated. Initial ensembles were generated as described

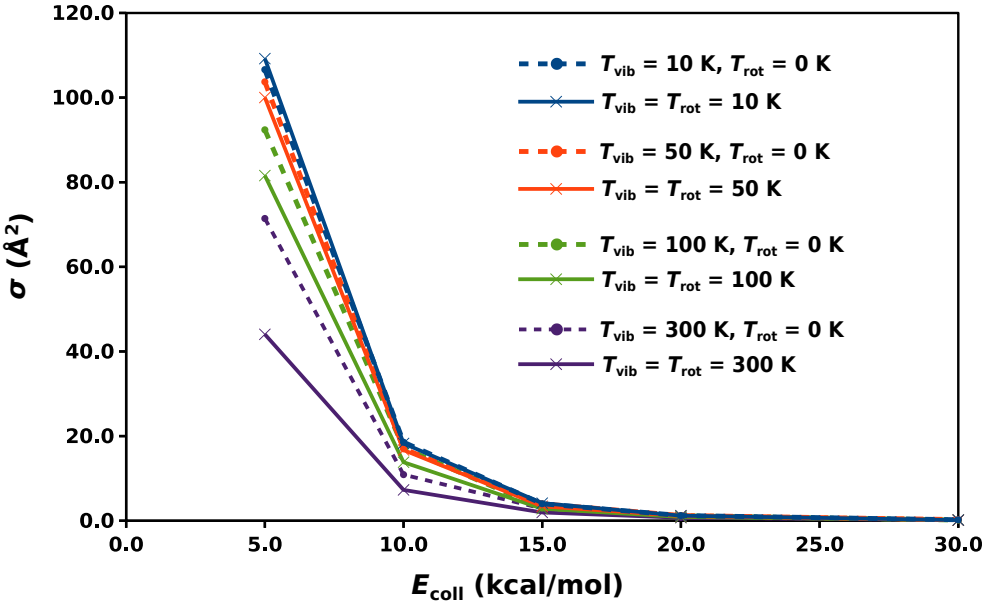


Figure 4.14: Variation of the cross section (σ) for the formation of the van der Waals complex in the entrance channel of the Diels-Alder reaction between dibromobutadiene and maleic anhydride as a function of the collision energy (E_{coll}) with different vibrational and rotational temperatures (T_{vib}, T_{rot}).

in section 4.2.1 and the impact parameter b was uniformly sampled in intervals of 1 Å until the maximum impact parameter was reached. For each set of initial conditions (E_{coll}, T_{vib}, T_{rot}, b) 500 trajectories were run for 10 ps. If at the end of a trajectory, the center of mass distance between the two molecules was < 15 Å, it was considered that a van der Waals complex had been formed. Figure 4.14 shows the cross section for the formation of complexes as a function of the collision energy. It can be seen that the cross section diminishes as the collision energy increases. In fact, the cross section is close to zero for $E_{\text{coll}} > 20$ kcal/mol.

The influence of vibrational and rotational temperature is also shown in figure 4.14. Increasing the energy of the system in either vibrational or rotational degrees of freedom reduces the cross section for complex formation. These findings can be explained by the stabilization of the well of the van der Waals complex of around 12 kcal/mol (figure 4.3). Thus when the system has collision energies above 15 kcal/mol or sufficiently high rotational and vibrational excitation, it can dissociate or not get trapped in the potential well at all.

4.3.4 Reactive collisions

For studying the Diels-Alder reaction itself, approximately 10^7 reactive molecular dynamics simulations were carried out. The sets of simulations that yield reactive events are summarized in table 4.1 and figure 4.15(a) displays trajectories of two reactive events. The vibrational temperature was 100 K for most of the trajectories in order to classically simulate vibrationally cold molecules as it is usually the case in collision experiments with supersonic molecular beams although some tests at higher vibrational temperatures have also been performed. Collision energies between 15 and 100 kcal/mol were sampled. The impact parameter was varied from 0 to 6 Å. The rotational temperature was 0, 1000, 2000 and 4000 K such that the influence of rotational excitation could be studied.

E_{coll} (kcal/mol)	b (Å)	T_{rot} (K)	T_{vib} (K)	# Reactive events	N
75	0 - 1	0	100	1	10^5
75	0	1000	100	1	10^5
75	0	2000	100	5	10^5
75	0	4000	100	4	10^5
100	1 - 2	0	100	1	10^5
100	0	1000	100	6	10^5
100	0	2000	100	8	10^5
100	0	4000	100	24	10^5
100	0	4000	100	174	$9{\cdot}10^5$
100	0 - 1	4000	100	129	10^6
100	1 - 2	4000	100	49	10^6
100	2 - 3	4000	100	39	10^6
100	3 - 4	4000	100	32	10^6
100	4 - 5	4000	100	3	10^6
15	0	8000	100	1	10^5
20	0	8000	100	2	10^5
75	0	0	1000	1	10^5
100	0	0	1000	2	10^5
Total				482	$7.1 \cdot 10^6$

Table 4.1: Initial conditions for the 482 recorded reactive events in terms of collision energy (E_{coll}), impact parameter (b) rotational temperature (T_{rot}) and vibrational temperature (T_{vib}). N is the total number of trajectories for each set of initial conditions.

In order to calculate a reactive cross section, we ran $5 \cdot 10^6$ trajectories for initial conditions $E_{\text{coll}} = 100$ kcal/mol, $T_{\text{rot}} = 4000$ K and $T_{\text{vib}} = 100$ K at which the largest number of reactive events were observed. The maximum impact parameter b_{max} was 5 Å which yields a cross section $\sigma = 2.13 \cdot 10^{-3}$ Å^2, corresponding to

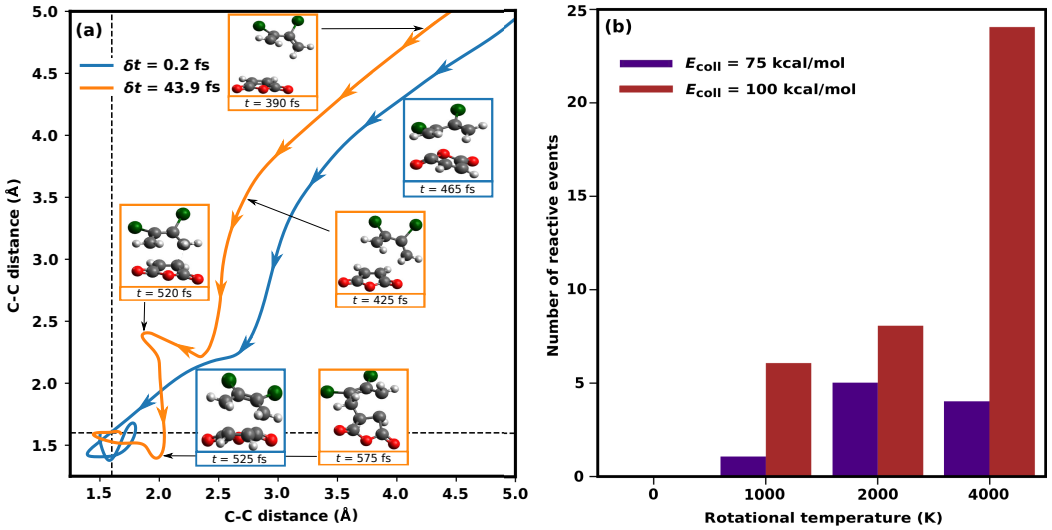


Figure 4.15: (a) Distances of the two C-C bonds formed along two reactive trajectories with different times δt elapsed between the formation of the first and the second bond. The blue and orange arrows indicate the direction of the trajectories. Some snapshots of structures along the trajectory are shown with their time stamps. The dashed, black lines at 1.6 Å indicate the geometrical threshold for a C-C bond formation. (b) Variation of the number of reactive events at collision energies 75 and 100 kcal/mol with vibrational temperature 100 K and impact parameter 0 Å as a function of the rotational temperature of the reactant molecules. 10^5 trajectories were run per collision energy and rotational temperature (see table 4.1).

a rate $k = 7.53 \cdot 10^{-14}$ cm^3 molecule^{-1} s^{-1}. A collision energy of 100 kcal/mol corresponds to a relative velocity of 3533 m/s which is at the very upper limit of what may be experimentally achievable [135].

As the MDP shows and figure 4.15(b) confirms, the rotational energy promotes the reaction, although reactive events are still rare (1 in 10^4). In order to further test this result, a rotational temperature of $T_{\text{rot}} = 8000$ K was used for collisions at $E_{\text{coll}} = 15$, 20 kcal/mol but few reactive events (1 - 2) were observed, indicating that collision energy is also needed for the reaction to take place. It should be noted that only 5 reactive events out of 482 are recorded without rotational excitation (see table 4.1).

With a vibrational temperature of 100 K, no reactive events could be observed with $E_{\text{coll}} < 75$ kcal/mol even though, in principle, with a collision energy of 10 kcal/mol the system would have enough energy to overcome the reaction barrier (figure 4.3). This might indicate that the reaction rate with such low E_{coll} is too small combined with the fact that at these energies the reaction could be partly complex mediated and take much longer times than simulated. For a

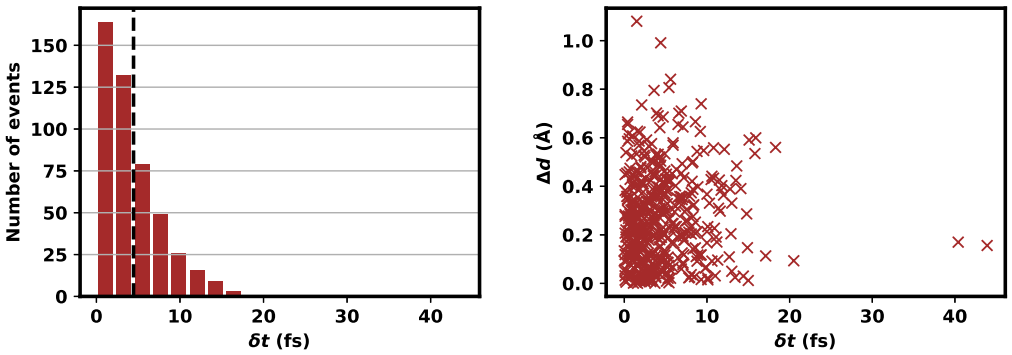


Figure 4.16: Left panel: histogram of the elapsed time (δt) for all reactive events. The mean of the distribution appears as a vertical line. Right panel: difference in the C-C distances of the two forming bonds at the transition state structure (Δd) versus elapsed time of formation of the two bonds (δt).

reactive event to take place, the molecules need to collide in a suitable relative orientation in order to overcome steric constraints and with the right distribution of energy, such that the TS can be reached. From figure 4.14, we know that there is no complex formation for $E_{\text{coll}} > 20$ kcal/mol and thus the reactions at the energies in figure 4.15 must be direct.

The influence of vibrational energy was tested by raising the vibrational temperature to $T_{\text{vib}} = 1000$ K and by individualy exciting some of the vibrational normal modes that appear to be active in the MDP. However, reactive events were so rare (0 - 2) that it can be concluded that vibrational activation of the reaction is weak, at least in this energy range.

The time δt elapsed between formation of the first and second bond was calculated for all the reactive events in order to determine the synchronicity of the reaction (figure 4.16). Out of 482 reactive events (see table 4.1), only two are slightly asynchronous with time lapses of 40.4 fs and 43.9 fs which are larger than the *ca.* 30 fs corresponding to a C-C bond vibrational period in cyclohexene [48, 51]. The difference between a synchronous and a slightly asynchronous process can be seen in figure 4.15(a). We have not found a correlation between the (a)symmetry of the TS structure in the dynamics and the (a)synchronicity of the process (figure 4.16) meaning that the elapsed time does not seem to depend on the (a)symmetry of the TS structure. Here, the TS geometry along the reactive trajectory is defined as the structure at the maximum in potential energy before the system crosses from the reactant surface to the product surface. When

recrossings occur, the last crossing point is taken for the definition of the TS structure.

4.4 Conclusions

The Diels-Alder reaction between MA and DBB has been studied using reactive molecular dynamics. The trajectories start with the two reactant molecules approaching each other as in a collision experiment. The minimum dynamic path indicates that rotations are important to drive the system towards the transition state. Furthermore, this finding has been confirmed by the fact that the majority of reactive collisions occur with rotational excitation. The presence of bromine substituents in the system accentuates the importance of rotations, but they were also found to be important for reactions of non substituted dienes. At the energies at which reactive events were observed, the cross section for the formation of van der Waals complexes in the entrance channel is almost zero and thus the reaction cannot be complex-mediated. Most of the observed reactive events are synchronous.

One of the fundamental aspects in reaction dynamics concerns the question which form of energy (translation, vibration or rotation) is most efficacious for the system to reach and surmount the transition state [141]. Back in the 1970s, when studying model atom plus diatom reactions, Polanyi formulated rules which relate the nature of the transition state (early or late) with the type of energy that promotes the reaction (translational or vibrational energy). Application [142] and generalization [140, 143] of these rules to polyatomic molecules remains a challenging undertaking, see for example the sudden vector projection (SVP) model [140, 143]. Analysis of a number of atom plus diatom (H+H_2, F+H_2, F+HCl) or atom plus triatom (H+H_2O, F+H_2O) reactions highlighted the cases under which rotations may play an important role in promoting or inhibiting a reaction [143]. The strength of the SVP model is that it requires only information about the normal modes and their directions in the reactant and transition state structures. On the other hand, the "sudden approximation" will not be applicable to situations in which internal vibrational relaxation (IVR) occurs or for collision energies much higher than the reaction threshold, as is the case in the present work. The present work suggests that rotations can play an important role for

reactions involving large excess of translational energy and the implications for reaction dynamics involving polyatomic molecules are exciting [25, 143–146].

As a next step, reactive molecular dynamics simulations for the cationic reaction between DBB and MA are performed in chapter 5. This reaction is expected to be faster and more complex since concerted and stepwise mechanisms are anticipated to coexist [135].

It would also be interesting to study this Diels-Alder reaction in solution which will provide deeper insights into molecular mechanisms governing solution-phase reactions [147]. Important additional effects concern the coupling between the solute and solvent degrees of freedom, potential barrier-lowering effects due to stabilization of the transition state, and changes in these properties depending on solvent identity. Recently, such molecular determinants have been studied for the Claisen rearrangement reaction using MS-ARMD [148]. It was found that both, electrostatic stabilization and localization of the solute contribute to lowering the barrier for the conversion of chorismate to prephenate in a protein environment. On the other hand, the rearrangement reaction for allyl vinyl ether did not show any barrier-lowering for the reaction in the protein. However, the reaction in water did show a barrier reduction because water can pack favorably around the solute and provide electrostatic stabilization of the transition state.

To the best of our knowledge, no simulation study had been performed in Diels-Alder reactions starting from the beginning of the reaction without steered dynamics. The present results indicate that the need of high collision energies together with rotational excitation and the low reaction rate of reaction will make the study of this reaction in single-collision experiments challenging.

Chapter 5

Reactive Molecular Dynamics of $[DBB + MA]^+$

5.1 Introduction

In this chapter, the gas-phase cationic Diels-Alder reaction between DBB and MA (see scheme 3.1) is studied by means of reactive molecular dynamics simulations. In chapter 4, the neutral Diels-Alder reaction was found to be synchronous, direct and promoted by rotational excitation of the reactant molecules. A similar study is performed here for the cationic variant of the same reaction. Again only the forward reaction, i.e. the formation of products, is studied. The trajectories start with the two reactant molecules approaching each other in order to simulate a collision experiment. Comparison with the results for the neutral reaction will provide insights into the differences on the mechanisms and dynamics of these two reactions.

The goal is to again have a model that describes the whole surface, even though the PES is more complex for the cationic variant, in order to study the dynamics of the reaction and answer questions such that whether the reaction is synchronous, whether it is complex-mediated and ways to promote it.

The parametrization of the model used for the simulations is explained in section 5.2 and its quality shown in section 5.3.1. The minimum dynamic path is analyzed in section 5.3.2, the van der Waals complex formation in the entrance channel is explored in section 5.3.3 and reactive trajectories are discussed in section 5.3.4.

5.2 Methods

Atomistic simulations were carried out with the CHARMM program [129] using multisurface reactive molecular dynamics (MS-ARMD) [99] based on an initial parametrization from SwissParam [130].

All bonds involving hydrogen atoms were flexible and the time step used in the simulations was $\Delta t = 0.1$ fs to ensure conservation of total energy. The velocity Verlet algorithm was used for the propagation of the equations of motion [114].

The initial SwissParam parametrization was modified in order to construct an MS-ARMD [99] force field for this reaction. Ensembles of structures for the parametrization of the MS-ARMD model were generated with CHARMM as explained in section 4.2.1. For the parametrization of the intermediate the final temperature was set to 100 K due to difficulties in getting low-energy structures. Besides, additional structures for this force field were generated through scans around the first new bond formed along the reaction. The force fields for the reactant, intermediate and product states were parametrized to reproduce DFT reference energies at the M06-2X/6-31G* level of theory in the same way that was previously explained in section 4.2.2. The force field used in chapter 4 for MA was also employed here. 1613 structures were used for the parametrization of the DBB^+ force field. The non-bonded terms of the reactant force field were parametrized with 2589 additional points. For the product and intermediate force fields, 2086 and 1785 structures were used, respectively. The product of the Diels-Alder reaction has two possible connectivities (see figure 4.2 where the bonds of the Diels-Alder product can be formed between atoms 6 and 12 and 7 and 11 or between atoms 6 and 11 and 7 and 12) and the intermediate has 4 different connectivities (depending on whether the first bond is formed between atoms 6 and 11, 6 and 12, 7 and 11 or 7 and 12). In order to make the parametrization

of the model permutation invariant, two and four different force fields for the description of the product and the intermediate were used, respectively.

For the collision simulations, ensembles of the individual molecules (MA and DBB^+) at 100 K of vibrational temperature were generated using CHARMM as described in section 4.2.1. The reactants were placed at an initial distance of 20 Å with a random relative orientation. The collision energy (E_{coll}) was tuned by varying the atomic velocities along the collision axis. The impact parameter (b) was uniformly sampled by displacing one of the molecules along a perpendicular axis (see figure 4.1). Rotational temperature (T_{rot}) was added following calculation of the moment of inertia tensor and assuming equipartition among the three rotational degrees of freedom [134].

The trajectories were considered reactive when they ended on the product force field of the MS-ARMD model. The reactive cross section σ was calculated as stated in equation (2.78) in chapter 2.

The total kinetic energy of the minimum-dynamic-path trajectories was decomposed as already explained in section 4.2.1.

5.3 Results and Discussion

5.3.1 Parametrization of the Reactive Force Fields

Figure 5.1 shows stationary points on the PES of the Diels-Alder reaction between DBB^+ and MA at the M06-2X/6-31G* level of theory (we will discuss the MS-ARMD energies later on). Contrary to what happened for the neutral reaction in chapter 4, s-*cis*-DBB^+ is a minimum on the PES. Due to the fact that both reactant molecules are symmetric, there are only two possible paths for the Diels-Alder reaction. These are referred to as "endo" and "exo" depending on the relative orientation of the reactants. Additionally, there is a path for the s-*trans*-DBB^+ to react. The endo product (P-endo$^+$) is taken as the zero of the energy scale in figure 5.1.

In order to study the dynamics of this system, we constructed a model PES that allows us to run a statistically significant number of trajectories, such that

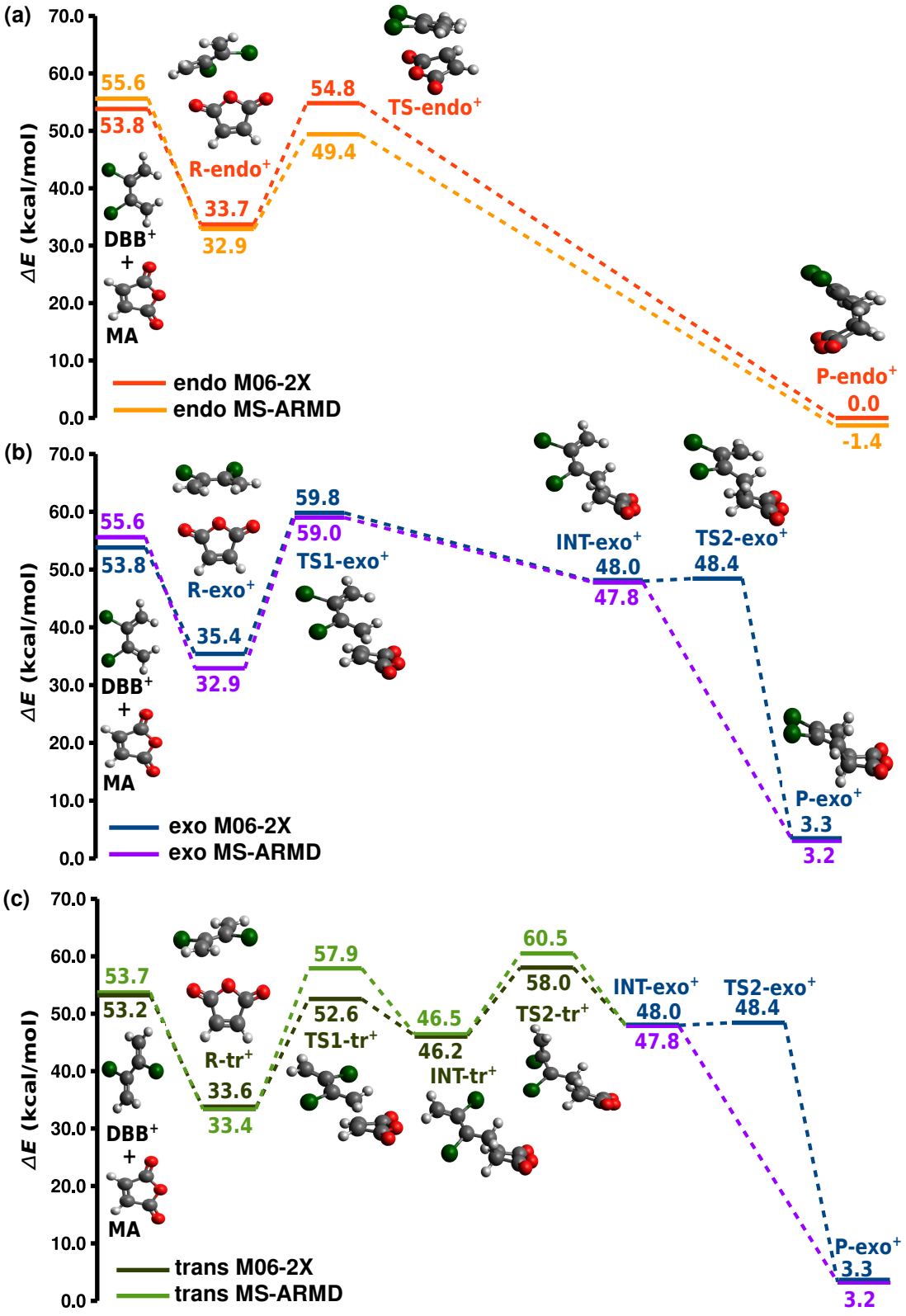


Figure 5.1: Potential energy surface for the three possible Diels-Alder reaction paths (a) endo, (b) exo and (c) trans between cationic dibromobutadiene (DBB$^+$) and maleic anhydride (MA) at the M06-2X/6-31G* and MS-ARMD levels of theory. The relative energies in kcal/mol with respect to the endo product (P-endo) as well as the structures of minima and transition states at the M06-2X level of theory are shown.

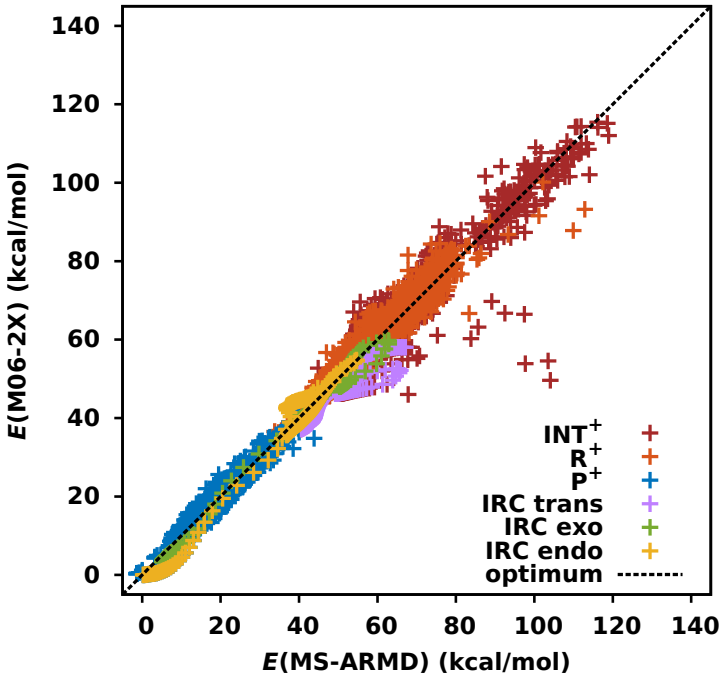


Figure 5.2: Energy correlation of 7055 M06-2X/6-31G* reference structures and the MS-ARMD model with a total RMSD of 2.9 kcal/mol.

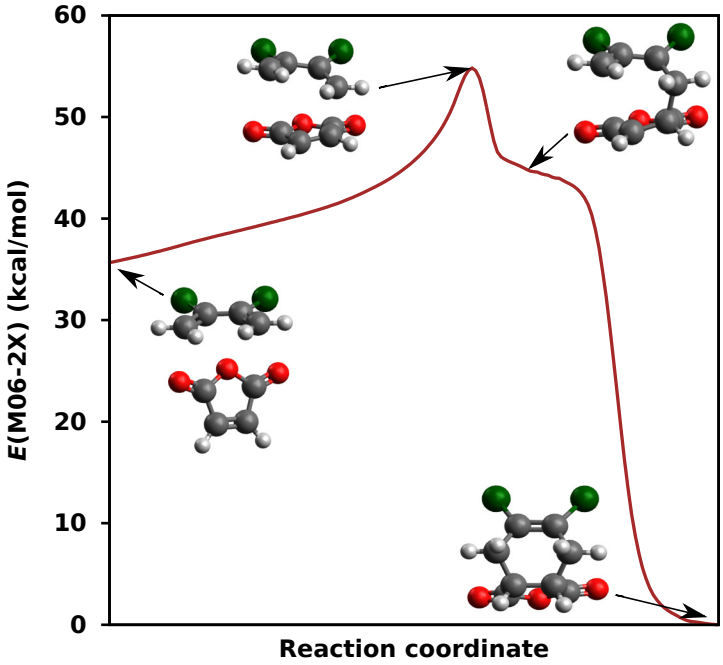


Figure 5.3: Endo IRC at the M06-2X/6-31G* level of theory. Some structures along the IRC are shown.

diverse initial conditions can be sampled. The quality of the parametrized MS-ARMD model is shown in figure 5.2. Points generated during the parametrization of the product force field (2086 structures), the intermediate (1785) and the parametrization of the non-bonded interactions of the reactant force field (2589) are shown, as well as the IRCs for the endo (169), the exo (192) and the trans (234) paths. In this case, the s-*trans* conformer of DBB^+ has been included and is a stable structure in the MS-ARMD model. The endo IRC was used for the parametrization of the GAPOs. It is important to take into account that the endo IRC is asymmetric since it has a plateau after the transition state (see figure 5.3). The structures in this region resemble those of the intermediate force field with one of the new C-C bonds formed. For this reason, the intermediate force field is active in this region which is an approximation because the endo path has no minimum there. However, this was the only way we could parametrize the surface such that a unique model could describe the three possible paths. The total RMSD is 2.9 kcal/mol over a range of 120 kcal/mol which we asses to be sufficient for a qualitative characterization of the dynamics of the system. There are some outliers in the intermediate force field (INT^+ in figure 5.2) but, since they have high energies in the MS-ARMD model, the system will rarely visit those regions and thus they should not play a big role in the dynamics.

In figure 5.1, a comparison between the potential energies for the different paths at the MS-ARMD and DFT levels of theory can be seen. The exo path (figure 5.1(b)) is well described although TS2-exo$^+$ does not exist in the MS-ARMD surface. As discussed above, the total RMSD of the surface is 2.9 kcal/mol and so it is reasonable that a TS that lies 0.4 kcal/mol above the minimum INT-exo$^+$ is not captured by our model. The MS-ARMD endo path (figure 5.1(a)) has its TS 5.4 kcal/mol lower than the reference energy which leads to the fact that the Diels-Alder reaction is more favorable, for MS-ARMD, than the dissociation of the van der Waals complex in the entrance channel while at the DFT level the dissociation and the reaction have similar energies. This mismatch will be translated into an overestimation of the rates of reaction through the endo path. Finally, in the trans path (figure 5.1(c)), TS1-tr$^+$ and TS2-tr$^+$ have an overestimation of their energies of 5.3 kcal/mol and 2.5 kcal/mol, respectively. The overestimation of the energy of the first TS should not be worrisome because the limiting step for the reactivity through this path is TS2-tr$^+$ which lays higher in energy than TS1-tr$^+$

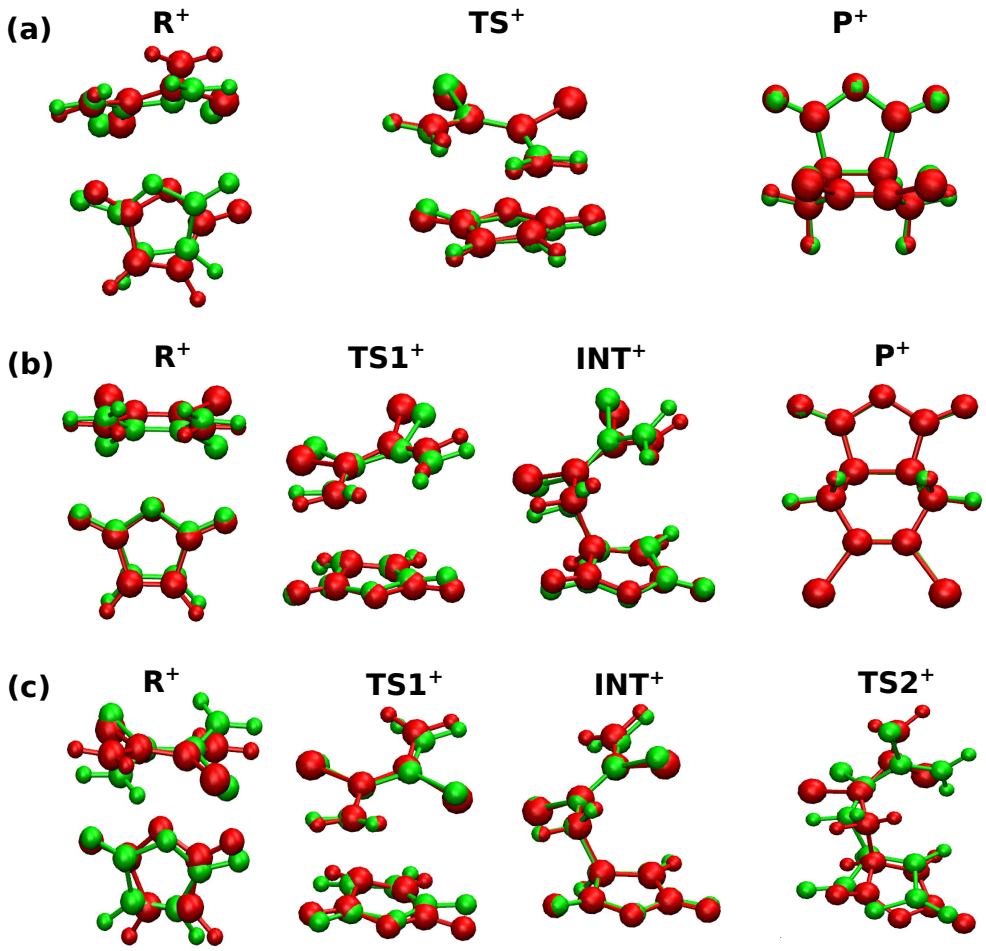


Figure 5.4: Comparison of optimized stationary-point structures along the (a) endo, (b) exo and (c) trans paths at DFT (red) and MS-ARMD (green) levels of theory.

also for MS-ARMD. However, the high energy of TS1-tr^+ will artificially elongate the lifetime of INT-tr^+.

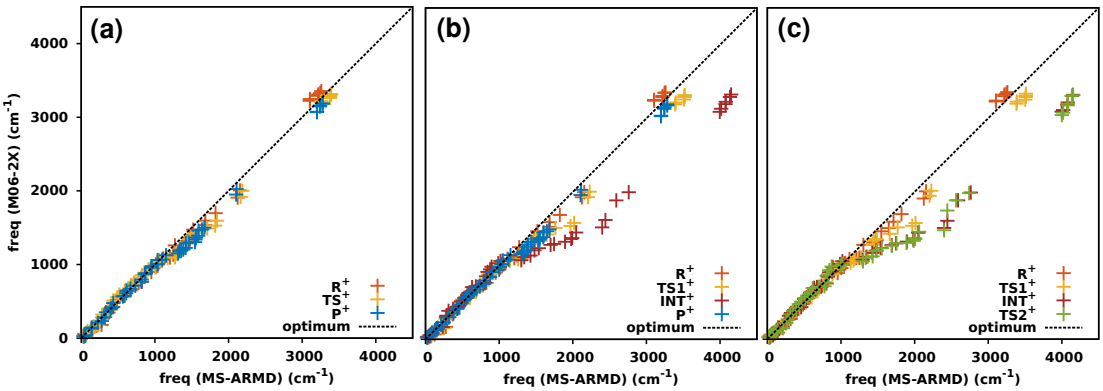


Figure 5.5: Comparison of harmonic frequencies at DFT and MS-ARMD levels of theory for the minimized structures of the (a) endo, (b) exo and (c) trans paths in figure 5.4.

For the validation of the MS-ARMD model the reactant, intermediate, transition state and product structures were minimized and compared to the DFT geometries in figure 5.4. The reactant structures of DFT and MS-ARMD have some differences that are expected since they are described with the same point charges like the dissociated structures and with isotropic van der Waals interactions which is an approximation. There are also some differences in the intermediate structures and TS2-tr^+ has exchanged positions for Br and CH_3, i.e. the cis/trans isomerization occurs in a different direction for MS-ARMD and for DFT. It is imporatnt to remember that the intermediate region of the surface is quite flat as it often happens in radical systems. The reference data points on this part of the surface are most certainly of lower quality than those in the reactant and the product region [135]. For this reason, special care is needed for the analysis of trajectories that spend time in the intermediate state region of the PES since the quality of the parametrization of this region is significantly worse than the reactant or product areas. This is due to the fact that the intermediate structure is much more flexible and it has a wider range of conformations than the reactant or the product. Since we want a unique force field for the s-*cis* intermediate (exo path) and the s-*trans* intermediate (trans path), the charges and equilibrium distances of bonds and angles are the same for both intermediates which also limits the quality of the model.

The harmonic frequencies have also been compared between DFT and the MS-ARMD model in figure 5.5 and again reflect the decreased quality of the intermediate force field as compared to the reactant and product force fields. The parametrized model is given in tables B.1-B.8 of appendix B.

5.3.2 Minimum Dynamic Path

The minimum dynamic path (MDP) [139] was calculated for the different paths as done for the endo path in chapter 4. The discussion will be centered around the exo path because it is the best one described by MS-ARMD. However, since the endo and trans paths are more favorable, the MDPs of these two paths will also be shown. The projection of the total kinetic energy along the MDP towards the reactants onto the degrees of freedom of DBB^+ and MA is shown in figure 5.6(a) as sums of translational, rotational and vibrational energies for the exo path.

At t = 0 fs, the system is at TS1-exo$^+$ and at t = 160 fs it has arrived at the reactant state. By projecting the total kinetic energy onto the different degrees of freedom of DBB$^+$ and MA, the active degrees of freedom can be identified. Figure 5.6(a) shows that vibrations contain the largest amount of energy for

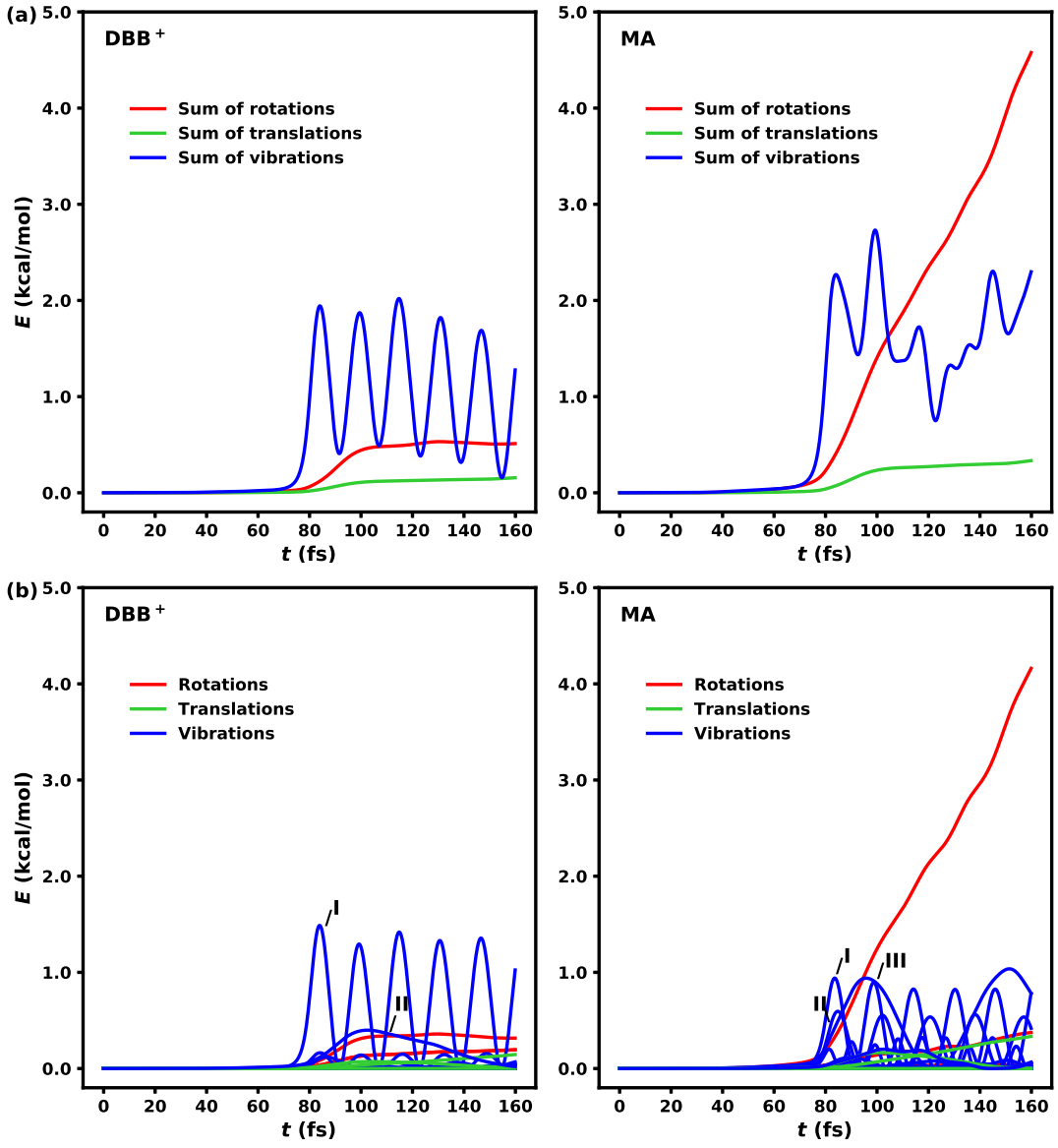


Figure 5.6: Projection of the total kinetic energy (E) onto the degrees of freedom of cationic dibromobutadiene (DBB$^+$) and maleic anhydride (MA) along the minimum dynamic path for the reaction of s-*cis*-DBB$^+$ + MA (a) summed into rotations, translations and vibrations (b) as individual traces. The trajectory starts at TS1-exo$^+$ and ends at the reactants. Most active vibrations for DBB$^+$: (I) out-of-plane symmetric bending of hydrogen and (II) skeleton out-of-plane asymmetric bend (cis/trans isomerization mode). For MA: (I) and (II) asymmetric and symmetric out-of-plane hydrogen bending, respectively and (III) asymmetric C=C out-of-plane bending.

DBB$^+$ while rotations contain the largest amount of energy for MA although vibrations are also active (see individual contributions in figure 5.6(b)). The rotational energy of DBB$^+$ and MA together accounts for 46% of the total kinetic energy while vibrational energy accounts for 48% and translational energy for only 6%. An identical result has been obtained from the direct decomposition of the total kinetic energy (figure 5.7). This is in clear contrast to the results of the neutral system where rotations accounted for 63% of the total kinetic energy and vibrations and translations for 19% and 18% respectively.

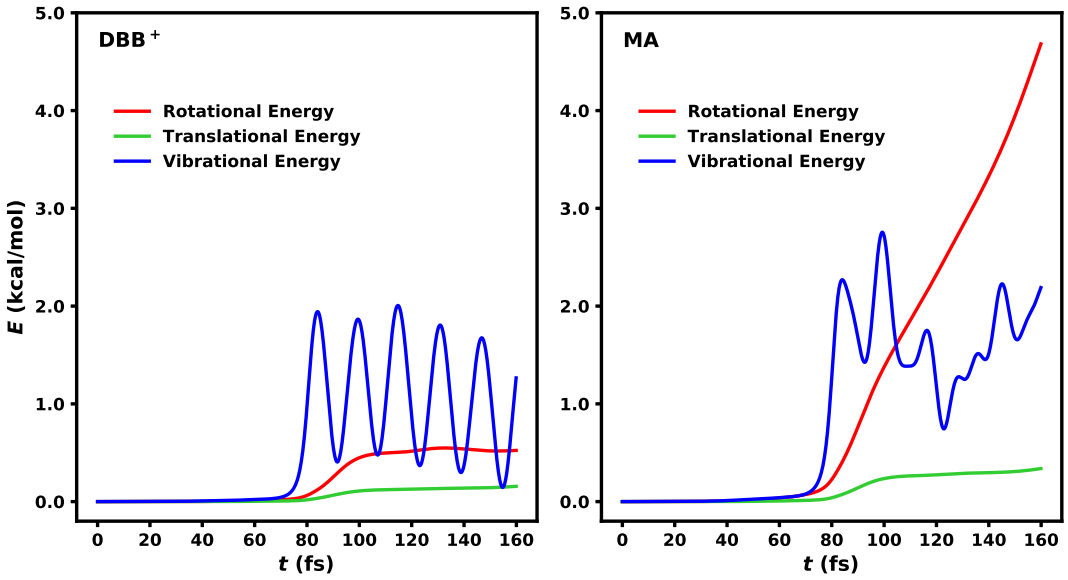


Figure 5.7: Decomposition of the total kinetic energy (E) along the minimum dynamic path into rotational, translational and vibrational energy as stated in section 4.2.1. The trajectory starts at TS1-exo$^+$ and ends at the reactants.

The excitation of vibrational modes could come from the asymmetry of the cationic TS that deforms the molecules more than the symmetric neutral TS does. The vibrations would then participate more in the reaction coordinate of the cationic system to bring the reactant molecules back to their equilibrium structures.

The projection of total kinetic energy onto the degrees of freedom of DBB$^+$ and MA along the endo and trans MDPs are shown in figures 5.8 and 5.9. For the endo path vibrational energy accounts for 54% of the total kinetic energy while rotational and translational energies account for 38% and 8%, respectively which indicates that rotational energy might be less important for the endo path than for the exo path. For the trans path the contributions of vibrational, rotational

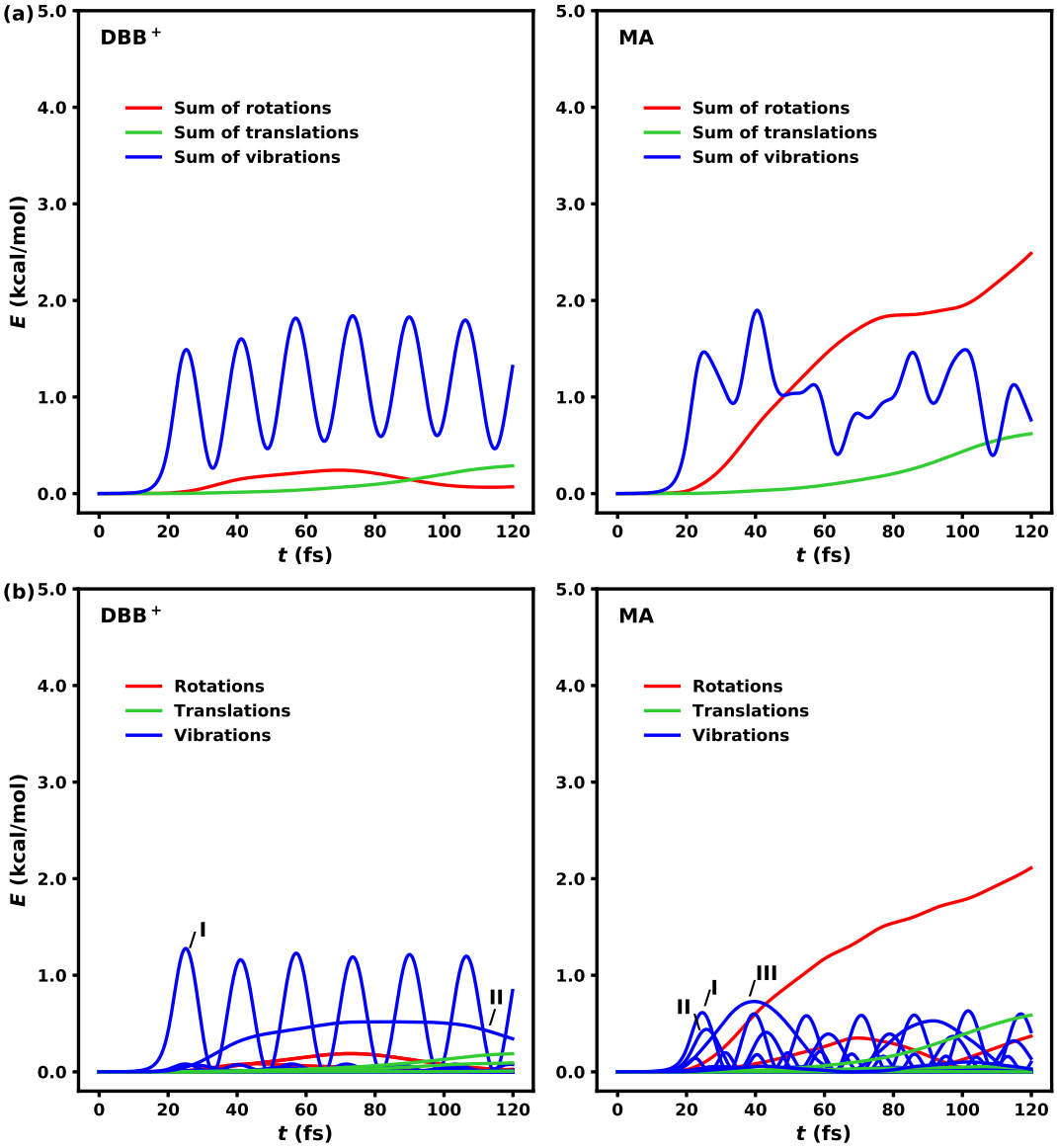


Figure 5.8: Projection of the total kinetic energy (E) onto the degrees of freedom of cationic dibromobutadiene (DBB$^+$) and maleic anhydride (MA) along the minimum dynamic path for the reaction of s-*cis*-DBB$^+$ + MA (a) summed into rotations, translations and vibrations (b) as individual traces. The trajectory starts at TS-endo$^+$ and ends at the reactants. Most active vibrations for DBB$^+$: (I) out-of-plane symmetric bending of hydrogen and (II) skeleton out-of-plane asymmetric bend (cis/trans isomerization mode). For MA: (I) and (II) asymmetric and symmetric out-of-plane hydrogen bending, respectively and (III) asymmetric C=C out-of-plane bending.

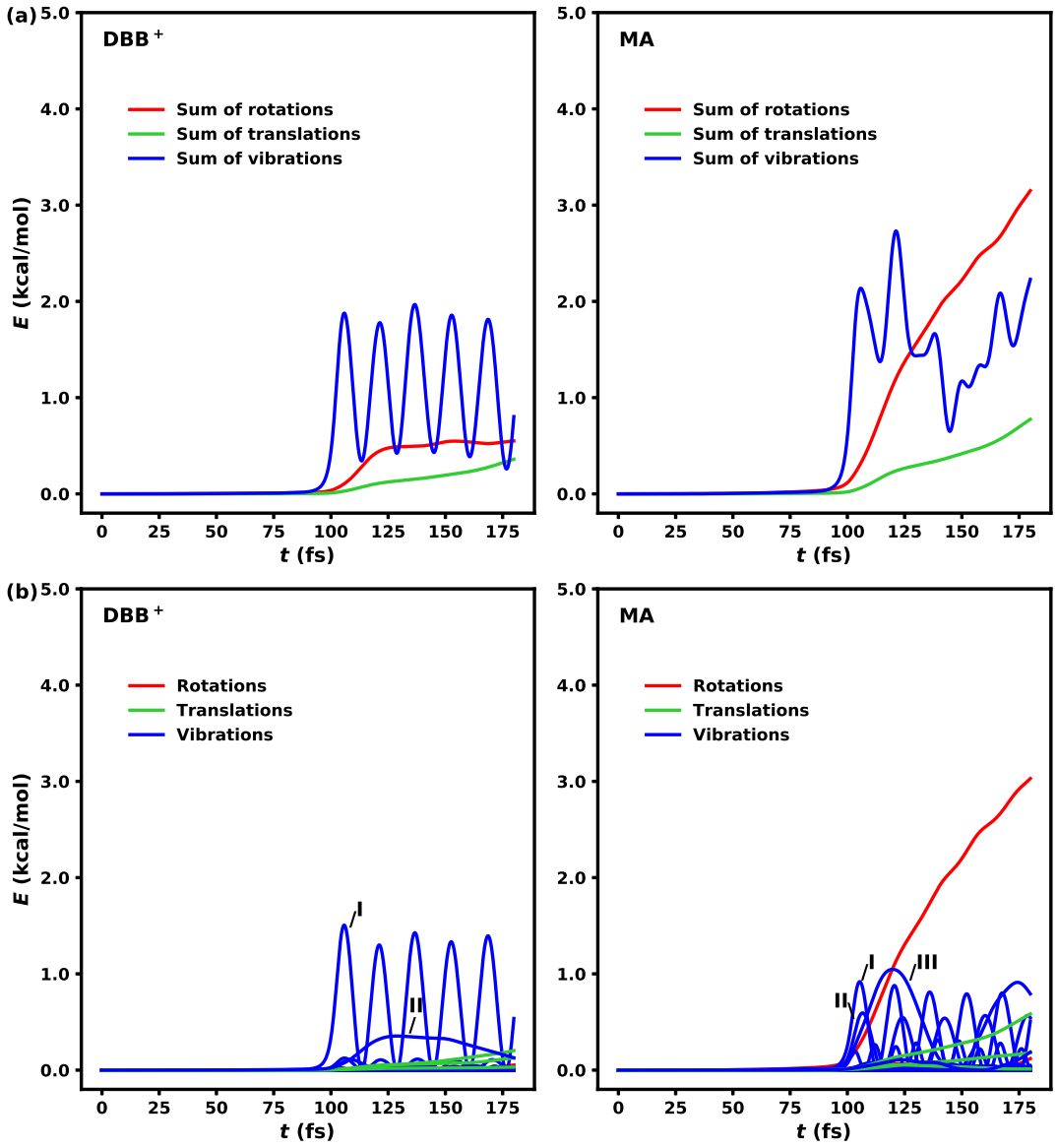


Figure 5.9: Projection of the total kinetic energy (E) onto the degrees of freedom of cationic dibromobutadiene (DBB$^+$) and maleic anhydride (MA) along the minimum dynamic path for the reaction of s-*trans*-DBB$^+$ + MA (a) summed into rotations, translations and vibrations (b) as individual traces. The trajectory starts at TS1-trans$^+$ and ends at the reactants. Most active vibrations for DBB$^+$: (I) out-of-plane symmetric bending of hydrogen and (II) skeleton out-of-plane symmetric bend. For MA: (I) and (II) asymmetric and symmetric out-of-plane hydrogen bending, respectively and (III) asymmetric C=C out-of-plane bending.

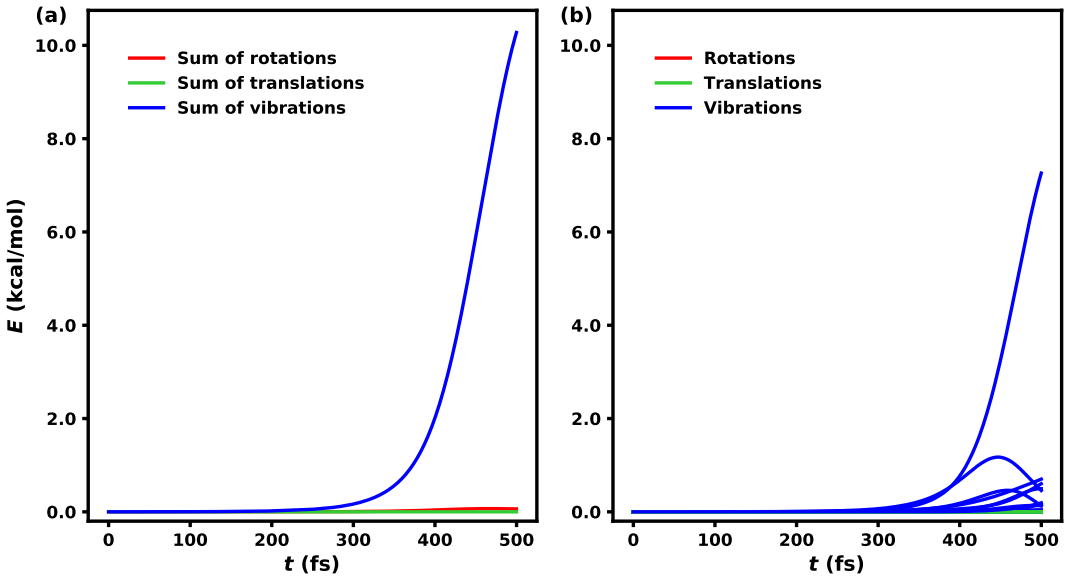


Figure 5.10: Projection of the total kinetic energy (E) onto the degrees of freedom of INT-tr$^+$ along the minimum dynamic path for the cis/trans isomerization of the intermediate (a) summed into rotations, translations and vibrations (b) as individual traces. The trajectory starts at TS2-trans$^+$ and ends at INT-tr$^+$. The most active vibration is the cis/trans isomerization mode.

and translational energy are 50%, 40% and 10%. The active vibrations of MA are the same in all paths. For DBB$^+$, they remain the same for the endo and exo trajectories and one mode changes for the trans path because the conformation of DBB$^+$ is different.

The MDP for the cis/trans isomerization of INT$^+$ has also been calculated. The total kinetic energy along this trajectory has been projected onto the degrees of freedom of INT-tr$^+$ and it is shown in figure 5.10. All of the energy goes into vibrational energy as expected for a unimolecular reaction. The most active vibration is the cis/trans isomerization mode and other low frequency skeleton vibrations are also active.

5.3.3 Cross section for the formation of van der Waals complexes in the entrance channel

The formation of van der Waals complexes in the entrance channel was studied in order to determine whether the reaction is direct (without the formation of

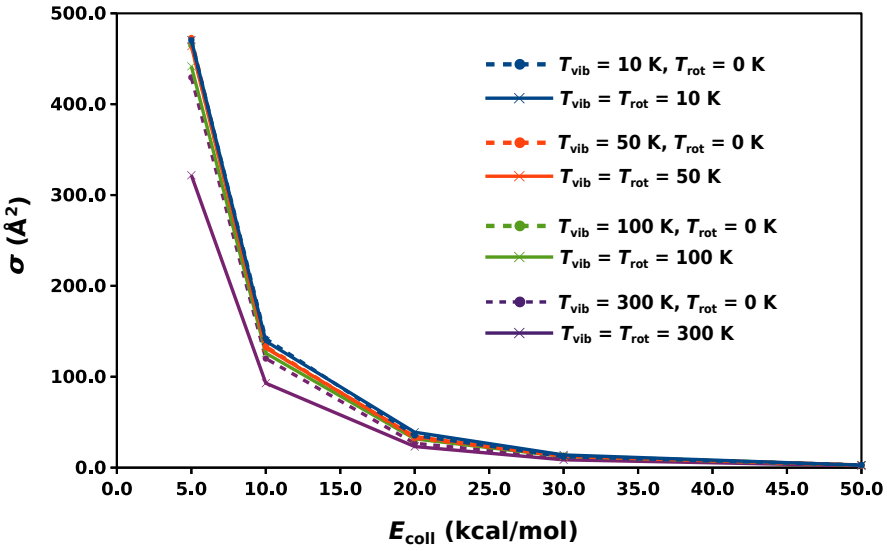


Figure 5.11: Variation of the cross section (σ) for the formation of the van der Waals complex in the entrance channel of the Diels-Alder reaction between cationic s-*cis*-dibromobutadiene and maleic anhydride as a function of the collision energy (E_{coll}) with different vibrational and rotational temperatures (T_{vib}, T_{rot}).

complexes) or complex-mediated. Initial ensembles were generated as described in section 4.2.1 and the impact parameter b was uniformly sampled in intervals of 1 Å until the maximum impact parameter was reached. For each set of initial

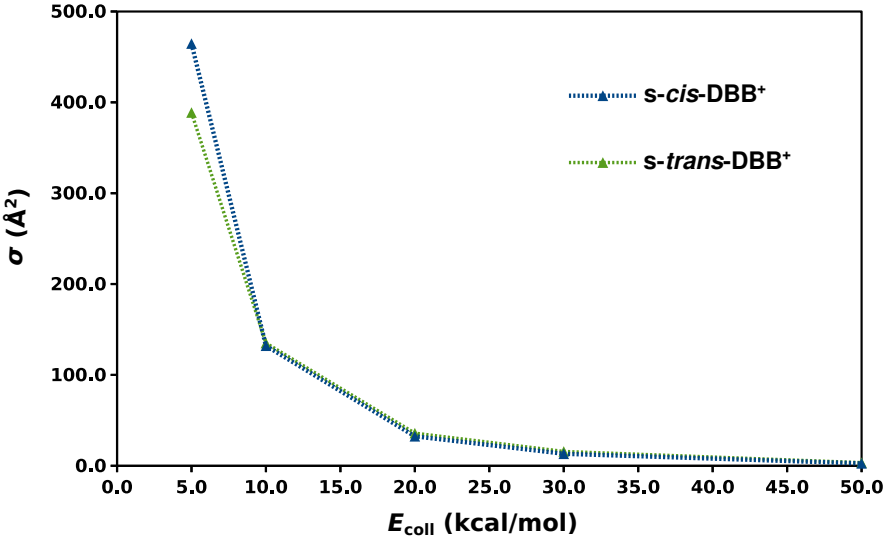


Figure 5.12: Variation of the cross section (σ) for the formation of the van der Waals complex in the entrance channel of the Diels-Alder reaction between cationic dibromobutadiene and maleic anhydride as a function of the collision energy (E_{coll}) for the s-*cis* and s-*trans* conformer of DBB^+ at $T_{\text{vib}} = 100$ K and $T_{\text{rot}} = 0$ K.

conditions (E_{coll}, T_{vib}, T_{rot}, b), 500 trajectories were run for 10 ps. If at the end of a trajectory, the center of mass distance between the two molecules was < 15 Å, it was considered that a van der Waals complex had been formed. Figure 5.11 shows the cross section for the formation of complexes as a function of the collision energy. It can be seen that the cross section diminishes as the collision energy increases. If we compare it to the neutral case, the maximum reached cross section is $\sigma \approx 475$ Å^2 while for the neutral case was $\sigma \approx 110$ Å^2. Besides, the cationic cross section decreases much slower and collision energies of 50 kcal/mol are needed to get $\sigma \approx 0$ Å^2. Rotational and vibrational energy have much less influence than for the neutral case. All these results reflect the fact that the cationic van der Waals complex minimum is ≈12 kcal/mol deeper than the neutral one.

The influence of the initial conformation of DBB^+ in the cross section for the formation of a van der Waals complex in the entrance channel is shown in figure 5.12. It can be seen that there is only a small difference at the lowest collision energies where the cross section for the *cis* conformer is higher than for the *trans* conformer due to the fact that $b_{max} = 16$ Å for s-*cis*-DBB^+ while it is 14 Å for s-*trans*-DBB^+. The difference in b_{max} could come from the difference in dipole moments of the two conformers.

5.3.4 Reactive collisions

For studying the Diels-Alder reaction itself, $1.8 \cdot 10^6$ reactive molecular dynamics simulations were carried out. The vibrational temperature was set to 100 K in order to classically simulate vibrationally cold molecules as it is usually the case in collision experiments with supersonic molecular beams and trapped ions. Collision energies of 50, 75 and 100 kcal/mol were sampled. The rotational temperature was 0, 2000 and 4000 K such that the influence of rotational excitation could be studied. The trajectories started with either s-*cis*-DBB^+ or s-*trans*-DBB^+ (see table 5.1). However, this does not necessarily mean that trajectories starting with the s-*trans* conformer follow the trans path of figure 5.1 since DBB^+ can isomerize upon collision with MA.

The trajectories were propagated until dissociation of the van der Waals complex, until product formation or up to a maximum simulated time of $t = 600$ ps. Due

E_{coll} (kcal/mol)	DBB^+	T_{rot} (K)	# Products	# Intermediates	# Complexes
50*	s-*cis*	0	1	19	13644
50*	s-*trans*	0	0	7	14828
50	s-*cis*	2000	24	14	829
50	s-*trans*	2000	7	25	884
50	s-*cis*	4000	57	34	33
50	s-*trans*	4000	29	40	50
75	s-*cis*	0	18	13	104
75	s-*trans*	0	4	3	174
75	s-*cis*	2000	43	20	5
75	s-*trans*	2000	22	15	17
75	s-*cis*	4000	106	19	5
75	s-*trans*	4000	69	29	23
100	s-*cis*	0	30	4	0
100	s-*trans*	0	6	0	0
100	s-*cis*	2000	92	0	0
100	s-*trans*	2000	47	3	2
100	s-*cis*	4000	153	0	0
100	s-*trans*	4000	65	0	0
Total	$1.8 \cdot 10^6$ trajectories		773	245	30598

Table 5.1: Initial conditions sampled for the recorded reactive events in terms of collision energy (E_{coll}), conformation of dibromobutadiene (DBB^+) and rotational temperature (T_{rot}) at a vibrational temperature $T_{vib} = 100$ K and impact parameter $b = 0$ Å. *All simulations were propagated until dissociation or until they reach the products up to a total time of $t = 600$ ps except for those with initial conditions $E_{coll} = 50$ kcal/mol, $T_{rot} = 0$ K that were only propagated until $t = 300$ ps.

to the long lifetime of the van der Waals complexes from the trajectories with initial conditions $E_{coll} = 50$ kcal/mol, $T_{rot} = 0$ K, these trajectories were only propagated for up to 300 ps (simulating 300 ps takes ~ 50 min of computational time).

Contrary to what happened for the neutral reaction in chapter 4, all sets of initial conditions but one (at $E_{coll} = 50$ kcal/mol, $T_{rot} = 0$ K for s-*trans*-DBB^+) yield products for the cationic reaction (see table 5.1). In this way it is clear that the cationic system is more reactive than its neutral counterpart (see section 4.3.4).

As the MDP suggested and figure 5.13 confirms, the rotational energy promotes the reaction even if the rotational degrees of freedom are less active for the cationic MDP than for the neutral case. In fact, there are almost 5 times more reactive

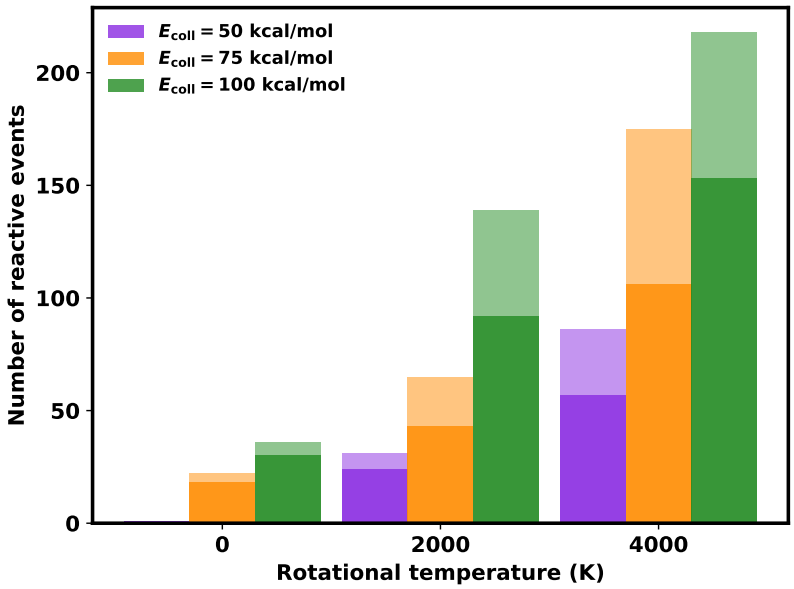


Figure 5.13: Stacked histogram for the variation of the number of reactive events at collision energies 50, 75 and 100 kcal/mol with vibrational temperature 100 K and impact parameter 0 Å as a function of the rotational temperature of the reactant molecules. Reactive events from trajectories that started with s-*cis* and s-*trans*-DBB$^+$ are represented in solid and transparent colors, respectively. 10^5 trajectories were run per collision energy, rotational temperature and initial conformation of DBB$^+$ (see table 5.1).

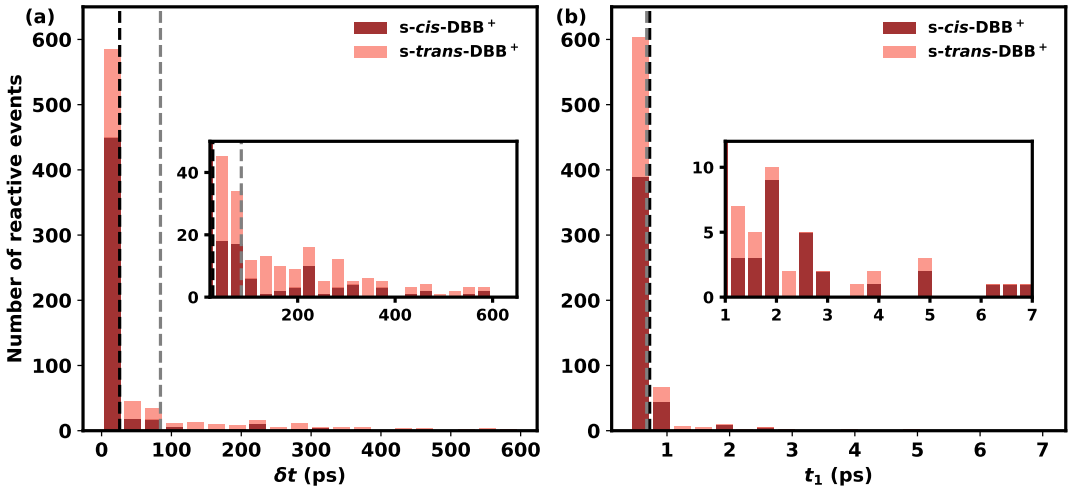


Figure 5.14: Stacked histogram of (a) the elapsed time (δt) and (b) the time of formation of the first bond (t_1) for all reactive events. Trajectories starting with s-*cis* and s-*trans*-DBB are in dark and light brown, respectively. The mean of the distributions appear as vertical lines in black for s-*cis*-DBB$^+$ and in gray for s-*trans*-DBB$^+$. The insets show magnifications of the tail of the distributions.

events at E_{coll} = 75 kcal/mol, T_{rot} = 4000 K than at E_{coll} = 100 kcal/mol, T_{rot} = 0 K even though they have similar total kinetic energies.

In order to determine the synchronicity of the reactions, the time δt elapsed between formation of the first and second bonds was calculated for all the reactive events as the time difference between the crossing time from the reactant force field to the intermediate force field (t_1) and the crossing time from the intermediate force field to the product force field (t_2) (see figure 5.14(a)). t_1 and t_2 are approximate time stamps for the formation of the first and second bonds since the system crosses surfaces at C-C distances longer than 1.6 Å that is the usual threshold for formation of these bonds. However, if the system stays in the second surface (the intermediate force field for t_1 and the product force field for t_2) we know that the bond is actually formed because otherwise the system would cross back to the first surface. Out of 773 reactive events (see table 5.1), only 59 are synchronous with $\delta t < 30$ fs [48, 51]. All the synchronous processes start from s-*cis*-DBB^+. The intermediate species has lifetimes on the order of picoseconds. It is important to remember that our model overestimates the lifetime of the intermediate due to the high activation barriers in the trans path of figure 5.1.

In order to determine whether the reaction is direct or complex-mediated, the time t_1 of surface crossing between the reactant and intermediate force fields is shown in figure 5.14(b) for all reactive trajectories. This time $t_1 < 2$ ps for the majority of trajectories and $t_1 < 7$ ps for all reactive trajectories indicating that they are direct events.

At the end of our simulated times there are still $\sim$ 30000 van der Waals complexes (see table 5.1) that could eventually form products. However, we have not recorded product formation with $t_1 > 7$ ps even though some of the van der Waals complexes live for 600 ps so this will be an unlikely and slow process. The 245 trajectories that are trapped in the intermediate region could eventually evolve to products or dissociate. As it can be seen in figure 5.14(a), the formation of the second bond in the tail of the distribution is again an unlikely event.

5.4 Conclusions

The Diels-Alder reaction between MA and DBB^+ has been studied using reactive molecular dynamics. The trajectories start with the two reactant molecules approaching each other as in a collision experiment. The minimum dynamic path indicates that rotations and vibrations are important to drive the system towards the transition state in clear contrast to the neutral case where vibrations did not seem to play an important role. This could come from the fact that the cationic transition state is asymmetric (the neutral was symmetric) while the reactant molecules are symmetric which leads to a greater deformation of the molecular structures that excites some vibrational modes along the MDP. Another difference to the neutral case is that the cationic system can form van der Waals complexes even at the high collision energies employed here. However, at the energies at which reactive events are recorded the reactions are direct and mostly asynchronous although synchronous processes can also take place. The cationic system is more reactive than the neutral one, as it can be seen from the comparison between figures 5.13 and 4.15(b). The maximum number of reactive events is obtained for a collision energy of 100 kcal/mol and a rotational temperature of 4000 K. Taking only those trajectories that start with the s-*cis* conformer of DBB for the cationic system, 153 reactions are recorded whereas only 24 are recorded for the same conditions in the neutral system. This was already expected from the difference in activation energies of the two systems in chapter 3.

It is important to keep in mind that the MS-ARMD model employed in this chapter is not well suited for trajectories that spend long times in the intermediate region since the quality of the model there is worse than in the reactant and product regions. Our model could be improved by including poorly described structures sampled by reactive trajectories in the parametrization of the intermediate force field. Another option would concern the parametrization of a different force field for each path. In the latter case, the model would not have a complete picture of the reaction but the quality of the individual paths would be significantly improved.

As a next step, it would be interesting to study the influence of vibrational energy for the reactive collisions and see whether it promotes the reaction as the minimum dynamic path seems to indicate. Additionally, since all the reactive trajectories performed in this work are head-on, running trajectories with different

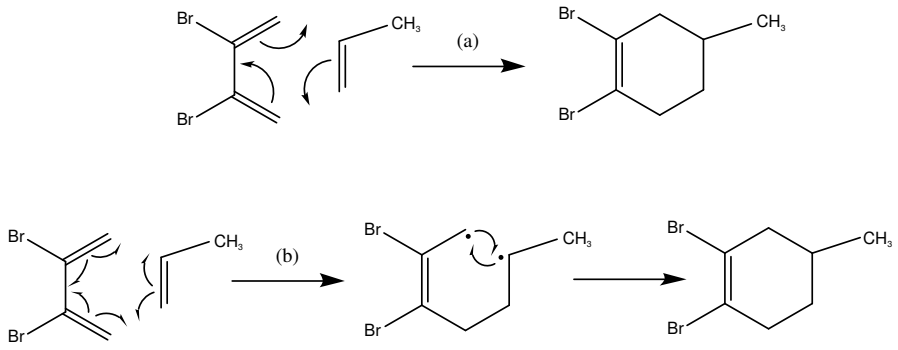
impact parameters would allow to study the influence of long range interactions. Another idea would be to extend the study to trajectories with low collision energies where most reactive events would take longer times than propagated here and would allow the exploration of differences between direct and complex-mediated mechanisms.

Chapter 6

Energetics of DBB + propene

6.1 Introduction

In this chapter, the neutral and cationic Diels-Alder reactions between 2,3-dibromo-1,3-butadiene (DBB) and propene (see scheme 6.1) are studied as an alternative to the reaction of DBB with maleic anhydride (MA). The neutral reaction is expected to have a higher activation energy than that of DBB + MA since propene does not have electron-withdrawing groups that would activate it as is the case for MA.



Scheme 6.1: Diels-Alder reaction between 2,3-dibromo-1,3-butadiene (DBB) and propene: (a) Concerted mechanism and (b) stepwise mechanism.

6.2 Methods

The potential energy surface (PES) of the system was calculated at the density functional theory (DFT) level of theory with the M06-2X functional [98] as recommended by Linder and Brinck [63] and using the Gaussian09 suite of codes [121]. The 6-31G* basis set was used as recommended in Refs. [62, 68]. All the geometry optimizations have been followed by frequency calculations and the connectivities of the transition states (TSs) with the minima have been checked by means of internal reaction coordinate (IRC) calculations [104].

6.3 Results

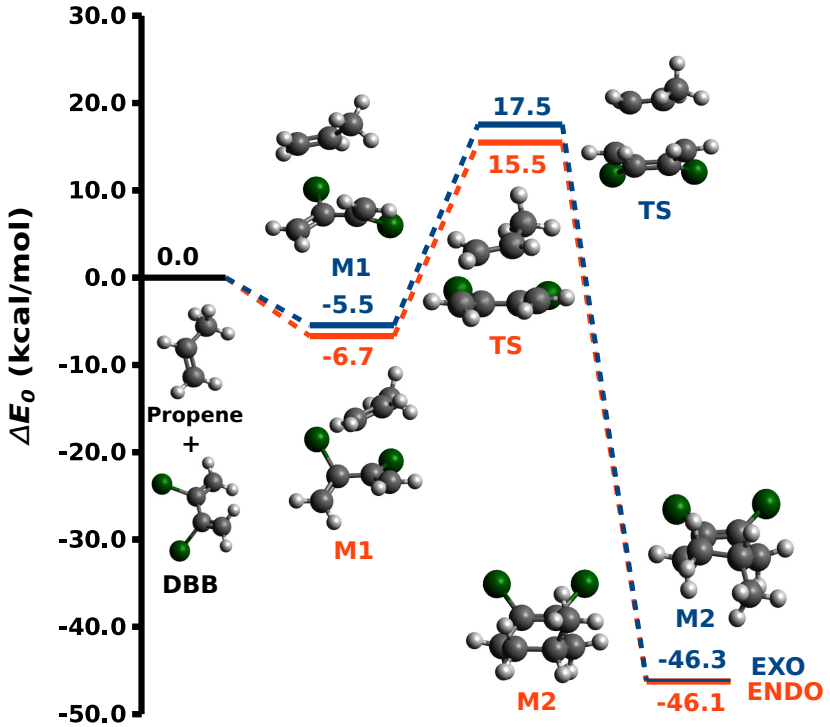


Figure 6.1: Potential energy surface for the two possible neutral Diels-Alder reaction paths (exo in blue and endo in orange) between 2,3-dibromo-1,3-butadiene (DBB) and propene at the M06-2X/6-31G* level of theory. The relative energies in kcal/mol with respect to the reactants as well as the structures of minima and transition states are shown. Green spheres represent bromine atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

Figure 6.1 depicts the stationary points found along the reaction coordinate of the neutral reaction which is concerted but asymmetric. At the TS, the C-C distances of the bonds that are formed are 2.23 Å and 2.32 Å. The reaction of DBB with MA has a symmetric TS with C-C distances of 2.27 Å. As it was expected, when using a non-activated diene, the Diels-Alder reaction has a higher activation energy. The energy barrier for the reaction with propene is 6.2 and 7 kcal/mol higher for the endo and exo paths, respectively than for the reaction with MA.

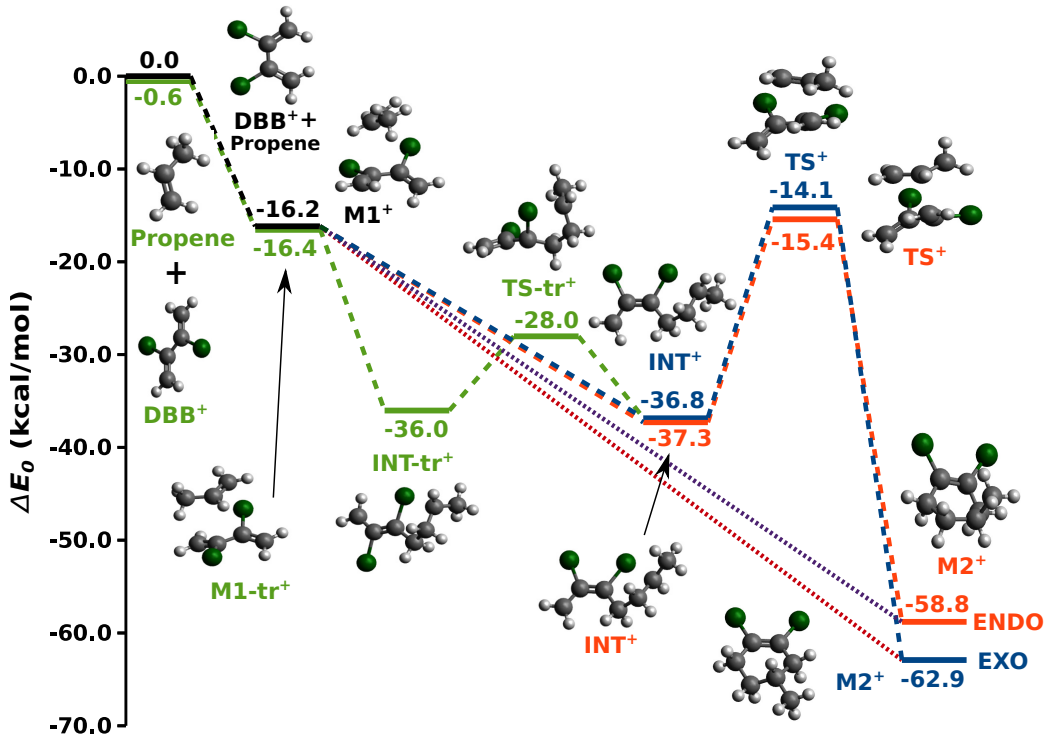


Figure 6.2: Potential energy surface for the three possible cationic Diels-Alder reaction paths (exo in blue, endo in orange and trans in green) between 2,3-dibromo-1,3-butadiene (DBB) and propene at the M06-2X/6-31G* level of theory. The finely dashed purple and red lines represent barrierless paths. The relative energies in kcal/mol with respect to the reactants as well as the structures of minima and transition states are shown. Green spheres represent bromine atoms, gray spheres carbon atoms and white spheres hydrogen atoms.

The cationic reaction is shown in figure 6.2. As it was the case for the reaction of DBB + MA in chapter 3, it can take place in a concerted or in a stepwise manner. The concerted mechanism is in this case a barrierless process. In the stepwise process the reaction occurs via an intermediate (INT^+) in which one of the C-C bonds has been formed. This first step is also barrierless. The second bond is then formed through TSs that lie lower than the reactants energy. Since

the reaction can be stepwise, the s-*trans* conformer of DBB may participate in the reaction. In this case the formation of the first bond is again barrierless but an additional TS appears for the *cis-trans* isomerization of the intermediate.

6.4 Conclusion

The neutral and cationic Diels-Alder reactions between dibromobutadiene and propene have been studied. The neutral reaction is concerted and asymmetric. Thus only the s-*cis* conformer of DBB is expected to participate in the reaction. On the contrary, the cationic reaction can be stepwise and all the stationary points found lie bellow the energy of the reactants which is common for cationic reactions.

When comparing it to the reaction of DBB + MA in chapter 3, the energy barriers for the neutral reaction with propene are higher than for the reaction with MA which can be explained by the fact that MA is an activated dienophile. For the cationic variant, the reaction with MA has energy barriers that lie above the reactants while the reaction of DBB^+ + propene has only submerged barriers or barrierless processes.

The study of the cationic reaction was done assuming DBB to be the cationic species since this is the ground state surface of the system. Another possibility would be that propene was the cationic species. Propene has an ionization potential higher than DBB and thus the system would start 10.5 kcal/mol higher than the zero of energy in figure 6.2. The question then would be whether an electron would hop to the propene before or after the reaction starts. In the first case, the reaction would take place on the PES represented in figure 6.2. In the second case excited states could be important to understand the course of the reaction.

Chapter 7

Energetics of $[N_2H + H_2O]^{+*}$

7.1 Introduction

Water exists in two forms that differ in the quantum number of the total nuclear spin I. *Para*-water has $I = 0$ while *ortho*-water has $I = 1$. These two forms of water are not interconvertible in isolated molecules. Hydrogen atoms have nuclear spin 1/2 which makes them fermions (the same as electrons). Hence, the total wave function must be antisymmetric for the exchange of the two hydrogen atoms of a water molecule. The electronic ground state of water has a symmetric electronic wave function. The wave function of the vibrational ground state is also symmetric. The nuclear wave function of a *para*-water molecule is antisymmetric and thus only rotational states with symmetric wave functions (states with even values of the sum of the asymmetric-top quantum numbers $K_a + K_c$) are accessible for this type of water molecules in the electronic-vibrational ground state. For *ortho*-water the opposite is true: in this case the nuclear wave function is symmetric and only antisymmetric rotational wave functions (states with odd $K_a + K_c$) are available such that the total wave function is again antisymmetric. For this reason, the ground state of *para*-water is the absolute rotational ground state while the ground state for *ortho*-water is the first excited rotational state.

*This chapter is based on the paper *Observation of different reactivities of para and ortho-water towards trapped diazenylium ions.* A. Kilaj, H. Gao, D. Rösch, U. Rivero, J. Küpper, S. Willitsch, Nature Communications 9 (2018) 2096.

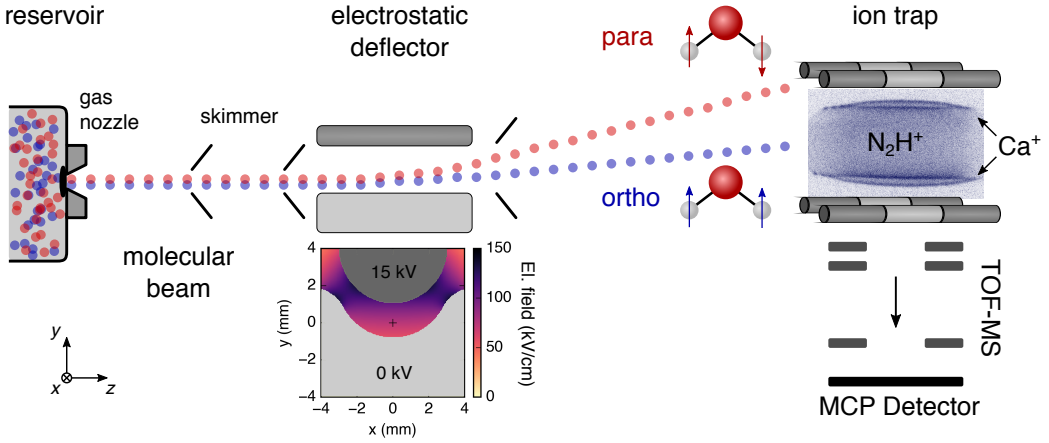


Figure 7.1: Schematic of the experimental setup. A pulsed molecular beam of water molecules seeded in argon emanates from a room-temperature reservoir through a pulsed gas nozzle and passes an electrostatic deflector. The inhomogeneous electric field inside the deflector (shown in the inset below) spatially separates *para*- and *ortho*-water molecules due to their different effective dipole moments. After the deflector, the beam is directed at an ion trap containing a Coulomb crystal of Ca^+ and sympathetically cooled N_2H^+ reactant ions (inset image). The products and kinetics of reactive collisions between N_2H^+ and H_2O are probed using a time-of-flight mass spectrometer (TOF-MS) [152].

Recent advances in experimental techniques allow the separation of species in a molecular beam by using inhomogeneous electrostatic fields as long as they have different effective dipole moments in the laboratory frame [85, 149]. The rotational states of a polar molecule differ in the effective dipole moment and are thus separable with this technique. Since the rotational state of water in its electronic-vibrational-ground state determines the total nuclear spin of the molecule, the separation of *ortho* and *para*-water can be achieved [150, 151].

In 2018, A. Kilaj *et al.* performed an experiment in which they observed the proton-transfer reaction of the two isomers of water with diazenylium (N_2H^+):

$$H_2O + N_2H^+ \rightarrow N_2 + H_3O^+. \tag{7.1}$$

The experimental setup is shown in figure 7.1. A molecular beam of internally cold water molecules is generated and passes through an electrostatic deflector that separates the *ortho*- from the *para*-water. In this way only one of the nuclear spin isomers hits the Coulomb crystal of Ca^+ ions inside which the sympathetically

cooled diazenylium molecular ions are trapped (in practice the method does not completely separate the two isomers but specific details can be found in Ref. [25]).

It was found that *para*-water reacts 23(9) % faster than *ortho*-water which can be explained in the framework of a rotationally adiabatic quantum capture theory for ion-molecule reactions provided that the process is barrierless [153, 154]. Within this framework, the difference in reactivity is a rotational effect induced by the fact that isomers with different nuclear-spin symmetry allow different rotational states. The *para*-water has, as a consequence, a smaller degree of rational averaging of the ion-dipole long-range interaction compared to the *ortho*-water.

In order to test whether the reaction is indeed barrierless, the potential energy surface (PES) of the reaction in equation (7.1) is studied in this chapter using *ab initio* methods.

7.2 Methods

The PES of the system has been calculated at the CCSD level of theory with the aug-cc-pVTZ basis set [87]. The geometry optimizations have been followed by frequency calculations. A relaxed scan was performed in which all the geometry parameters were optimized except for a distance that was constrained at a given value and used as an approximate representation of the reaction coordinate. All the calculations have been performed with Gaussian09 [121].

7.3 Results

The reaction (7.1) was studied at the CCSD/aug-cc-pVTZ level of theory and stationary points along the PES are shown in figure 7.2. This reaction appears to be barrierless and to proceed via formation of a pre-reaction complex in which one hydrogen atom is shared between the nitrogen and water moieties. The energy of the products was found to be 47.4 kcal/mol lower than that of the reactants. In order to test whether the reaction is indeed barrierless, electronic structure calculations were carried out. We assumed for this purpose that the O-H^+ distance approximates the reaction coordinate. We constrained this distance

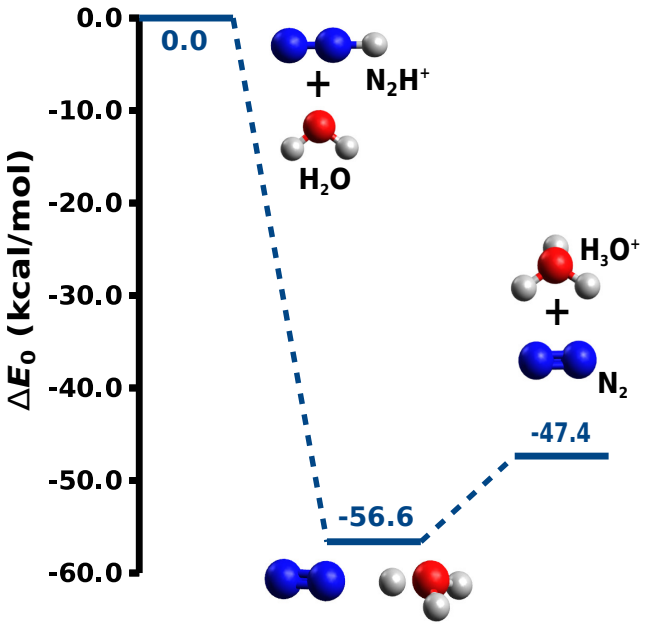


Figure 7.2: Potential-energy surface for the proton transfer reaction between N_2H^+ and H_2O at the CCSD/aug-cc-pVTZ level of theory. The relative energies with respect to the reactants as well as the structures of the stationary points are shown. Blue, red and white spheres represent nitrogen, oxygen, and hydrogen atoms.

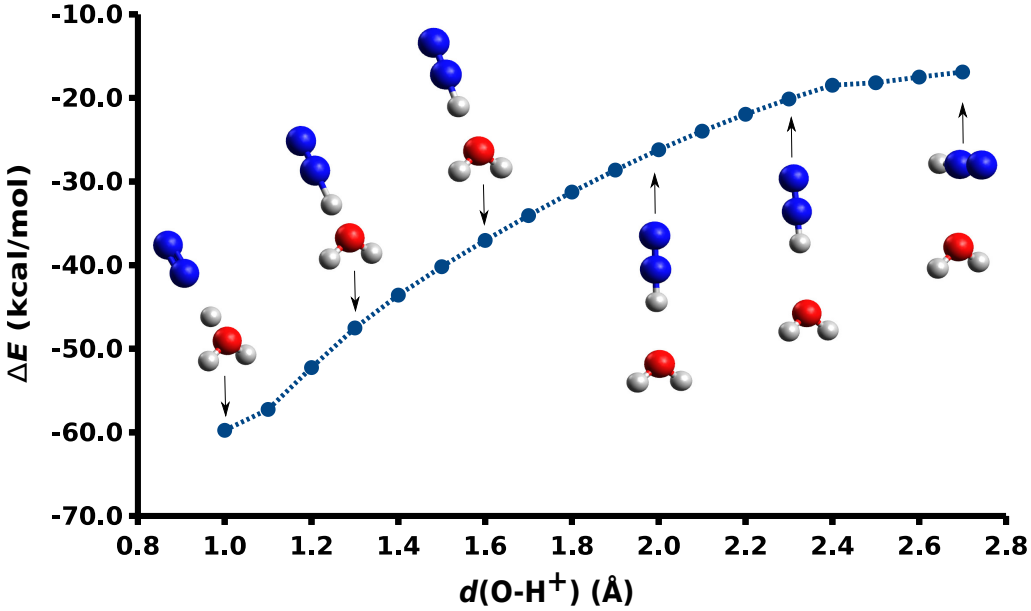


Figure 7.3: Optimized structures along the reaction coordinate of the proton transfer between N_2H^+ and H_2O at the CCSD/aug-cc-pVTZ level of theory. The energy is relative to that of the reactants. Blue spheres represent nitrogen atoms, red spheres oxygen atoms and white spheres hydrogen atoms.

while optimizing all other internal coordinates at the CCSD/aug-cc-pVTZ level of theory. The resulting electronic energies for the optimized structures along the O-H^+ coordinate are shown in figure 7.3. No transition state could be located within O-H^+ distances ranging from 1 to 2.6 Å since the energy continuously decreases from the reactants towards the products. These results indicate that the reaction is barrierless and vindicate the application of capture theory to model the reaction kinetics.

7.4 Conclusion

The proton-transfer reaction between N_2H^+ and H_2O has been studied at the CCSD/auc-cc-pVTZ level of theory and found to be barrierless as the experiment suggested.

This work is an example of how theory can help understand experiments. As it happened in chapters 4 and 5, the rotational degrees of freedom of the reactant molecules play here a key role in the rate of reaction.

Chapter 8

Conclusions and Outlook

The main goal of this work was to support and guide experiments through computational studies. In order to achieve this, three systems have been explored taking into consideration the experimental conditions at which they would be performed and trying to obtain theoretical results that can be experimentally studied.

An in-depth computational study of the neutral and cationic Diels-Alder reactions of 2,3-dibromo-1,3-butadiene (DBB) with maleic anhydride (MA) has been performed as a first step for a combined experimental and theoretical study. Density functional theory calculations predict the neutral reaction to be concerted with a symmetric transition state and thus only the s-*cis* conformer of the diene can yield the Diels-Alder product. Reactive atomistic simulations using multi-surface adiabatic reactive molecular dynamics (MS-ARMD) and the PhysNet neural network indicate that the reaction is concerted and that rotational excitation and high collision energies are needed to record reactive events. Comparison with the reaction between butadiene and MA shows that the bromine substituents of DBB accentuate the importance of molecular rotations. Analysis of the possible isomerizations and fragmentations of the Diels-Alder product show that the most possible decomposition of the product is to go back to reactants.

The cationic reaction, on the other hand, is predicted to have a concerted and a stepwise pathway both with asymmetric transition states. The s-*cis* and s-*trans* conformers of DBB^+ could participate although the reaction with s-*cis*-DBB^+ has

lower activation barriers and thus should be faster. Reactive atomistic simulations using MS-ARMD indicate that the reaction is direct, mostly asynchronous and can be promoted with rotational excitation of the reactant molecules. It would be interesting to study whether vibrational energy could also promote it as the minimum dynamic path seems to indicate. Besides, simulations at low collision energies would allow to study complex-mediated reactive events. In this work, DBB was the cationic species although, experimentally, MA would be the cationic; such that a beam of conformationally selected DBB would collide with sympathetically cooled MA. The system with MA^+ is an excited state since the ionization potential of DBB is ∼45 kcal/mol lower than that of MA. Studies on the charge transfer from MA^+ to DBB as the two species approach each other would be helpful to understand future experimental results. A different experiment could start with cationic DBB and seed the molecular beam with neutral MA. This would also allow to probe a wider range of collision energies.

Additionally, the neutral and cationic Diels-Alder reactions between DBB and propene have been studied by means of DFT theory. The neutral reaction is predicted to be concerted with asymmetric transition states and higher activation energies than the reaction with MA. The cationic reaction can be concerted or stepwise with submerged barriers and lower activation energies than the reaction with MA.

Finally, the reaction between nuclear-spin-selected water and sympathetically cooled N_2H^+ has been studied in order to aid with the interpretation of experimental results and found to be barrierless.

This thesis is an example of how experiments can benefit from theoretical studies.

The neutral Diels-Alder reaction is predicted to be experimentally very challenging to study due to the small reactive cross sections and the need for rotational excitation. On the other hand, the cationic reaction is much faster and the experiments on these systems benefit from the fact that one of the reactants can be trapped resulting in longer interaction times of the reactants than those of crossed molecular beam experiments.

Molecular dynamics studies on small molecules are crucial to understand experimental results and would benefit from the development of faster methods to parametrize accurate models for the reactions of interest. The bottleneck in MS-ARMD is the parametrization of the force-fields describing the different states of

the system. Better optimization algorithms would result on faster and more usable parametrization schemes. Another problem of MS-ARMD is the description of the transition state regions by mixing the different force fields in combination with Gaussian times polynomials functions. The problem with this description is that if the reactant and product states are symmetric, the model at the transition state region will not be able to accurately describe asymmetric transition states. Neural-network approaches could potentially be used in this region which would allow to have more flexibility without the computational cost and possible problems of using a neural-network description for the whole surface. Since a force field is defined from the concept of a bond, it is ensured that the molecule will stay bonded while a neural network can have unphysical behaviors in poorly sampled regions of the potential energy surface. Moreover, point charges force fields, with its limitations, have a straight forward way of including charged systems. In neural networks built from "amons", the inclusion of charged system is much complexer and still in development.

Bibliography

[1] M. Karplus, Angew. Chem. Int. Ed. **53**, 9992 (2014).

[2] M. Karplus, R. N. Porter, and R. D. Sharma, J. Chem. Phys. **40**, 2033 (1964).

[3] M. Karplus, R. N. Porter, and R. D. Sharma, J. Chem. Phys. **43**, 3259 (1965).

[4] M. Karplus, R. N. Porter, and R. D. Sharma, J. Chem. Phys. **45**, 3871 (1966).

[5] M. Meuwly, WIREs Comput. Mol. Sci. **9**, e1386 (2019).

[6] F. Trovato and G. Fumagalli, Biophys. Rev. **9**, 941 (2017).

[7] L. Dunoyer, C. R. Hebd. Séances Acad. Sci. **152**, 592 (1911).

[8] W. Gaede, Phys. Zeit. **6**, 758 (1905).

[9] G. Meyer, Verh. d. deutsch Phys. Ges. **10**, 753 (1907).

[10] D. Hershbach, Molecular Dynamics of Elementary Chemical Reactions, in *Nobel Lecutres, Chemistry 1981-1990*, edited by B. G. Malmström, p. 265, World Scientific Publishing, Singapore, 1992.

[11] A. Osterwalder, EPJ Techniques and Instrumentation **2**, 10 (2015).

[12] P. Casavecchia, Rep. Prog. Phys. **63**, 355 (2000).

[13] R. T. Skodje, D. Skouteris, D. E. Manolopoulos, S.-H. Lee, F. Dong, and K. Liu, J. Chem. Phys. **112**, 4536 (2000).

[14] J. Häger, S. Zech, D. Glatzer, M. Fink, and H. Walther, Rev. Sci. Instrum. **65**, 407 (1994).

[15] M. Kirste, X. Wang, H. C. Schewe, G. Meijer, K. Liu, A. van der Avoird, L. M. C. Janssen, K. B. Gubbels, G. C. Groenenboom, and S. Y. T. van de Meerakker, Science **338**, 1060 (2012).

[16] W. D. Phillips and H. Metcalf, Phys. Rev. Lett. **48**, 596 (1982).

[17] S. D. Hogan, D. Sprecher, M. Andrist, N. Vanhaecke, and F. Merkt, Phys. Rev. A **76**, 023412 (2007).

[18] E. Vliegen, S. D. Hogan, H. Schmutz, and F. Merkt, Phys. Rev. A **76**, 023405 (2007).

[19] A. T. J. B. Eppink and D. H. Parker, Rev. Sci. Instrum. **68**, 3477 (1997).

[20] L. Holmegaard, J. H. Nielsen, I. Nevo, and H. Stapelfeldt, Phys. Rev. Lett. **102**, 023001 (2009).

[21] F. Filsinger, J. Küpper, G. Meijer, J. L. Hansen, J. Maurer, J. H. Nielsen, L. Holmegaard, and H. Stapelfeldt, Angew. Chem. Int. Ed. **48**, 6900 (2009).

[22] F. Filsinger, G. Meijer, H. Stapelfeldt, H. N. Chapman, and J. Küpper, Phys. Chem. Chem. Phys. **13**, 2076 (2011).

[23] S. Y. T. van de Meerakker, H. L. Bethlem, N. Vanhaecke, and G. Meijer, Chem. Rev. **112**, 4828 (2012).

[24] J. H. Nielsen, P. Simesen, C. Z. Bisgaard, H. Stapelfeldt, F. Filsinger, B. Friedrich, G. Meijere, and J. Küpper, Phys. Chem. Chem. Phys. **13**, 18971 (2011).

[25] A. Kilaj, H. Gao, D. Rösch, U. Rivero, J. Küpper, and S. Willitsch, Nature Comm. **9**, 2096 (2018).

[26] Y.-P. Chang, K. Dlugolecki, J. Küpper, D. Rösch, D. Wild, and S. Willitsch, Science **342**, 98 (2013).

[27] E. L. Raab, M. Prentiss, A. Cable, S. Chu, and D. E. Pritchard, Phys. Rev. Lett. **59**, 2631 (1987).

[28] W. Paul, Rev. Mod. Phys. **62**, 531 (1990).

[29] C. J. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases* (Cambridge University Press, 2001).

[30] M. R. Tarbutt, Contemporary Physics **59**, 356 (2018).

[31] O. Dulieu and S. Willitsch, Europhysics News **48**, 17 (2017).

[32] W. W. Smith, O. P. Makarov, and J. Lin, J. Mod. Opt. **52**, 2253 (2005).

[33] F. H. J. Hall, P. Eberle, G. Hegi, M. Raoult, M. Aymar, O. Dulieu, and S. Willitsch, Mol. Phys. **111**, 2020 (2013).

[34] F. H. J. Hall, M. Aymar, M. Raoult, O. Dulieu, and S. Willitsch, Mol. Phys. **111**, 1683 (2013).

[35] Z. Meir, T. Sikorsky, R. Ben-shlomi, N. Akerman, Y. Dallal, and R. Ozeri, arXiv preprint arXiv:1603.01810 (2016).

[36] K. Liu, Annu. Rev. Phys. Chem. **67**, 91 (2016).

[37] H. C. Hsu, M.-T. Tsai, Y. A. Dyakov, and C.-K. Ni, Int. Rev. Phys. Chem. **31**, 201 (2012).

[38] J. P. Gordon, H. J. Zeiger, and C. H. Townes, Phys. Rev. **99**, 1264 (1955).

[39] J. Karczmarek, J. Wright, P. Corkum, and M. Ivanov, Phys. Rev. Lett. **82**, 3420 (1999).

[40] D. M. Villeneuve, S. A. Aseyev, P. Dietrich, M. Spanner, M. Y. Ivanov, and P. B. Corkum, Phys. Rev. Lett. **85**, 542 (2000).

[41] F. Rosca-Pruna and M. J. J. Vrakking, Phys. Rev. Lett. **87**, 153902 (2001).

[42] W. H. al-Qady, R. C. Forrey, B. H. Yang, P. C. Stancil, and N. Balakrishnan, Phys. Rev. A **84**, 054701 (2011).

[43] A. S. Meijer, Y. Zhang, D. H. Parker, and W. J. van der Zande, Phys. Rev. A **76**, 023411 (2007).

[44] O. Diels and K. Alder, Justus Liebigs Ann. Chem. **460**, 98 (1928).

[45] K. Ishihara and A. Sakakura, *Comprehensive Organic Synthesis*, 2 ed. (Elsevier, Oxford, 2014).

[46] K. N. Houk, J. González, and Y. Li, Acc. Chem. Res. **28**, 81 (1995).

[47] D. Yepes, O. Donoso-Tauda, P. Pérez, J. S. Murray, P. Politzer, and P. Jaque, Phys. Chem. Chem. Phys. **15**, 7311 (2013).

[48] M. A. F. de Souza, E. Ventura, S. A. do Monte, J. M. Riveros, and R. L. Longo, J. Comput. Chem. **37**, 701 (2016).

[49] L. R. Domingo, J. Chil. Chem. Soc. **2615**, 59 (2014).

[50] B. A. Horn, J. L. Herek, and A. H. Zewail, J. Am. Chem. Soc. **118**, 8755 (1996).

[51] E. W.-G. Diau, S. De Feyter, and A. H. Zewail, Chem. Phys. Lett. **304**, 134 (1999).

[52] N. J. Saettel, O. Wiest, D. A. Singleton, and M. P. Meyer, J. Am. Chem. Soc. **124**, 11552 (2002).

[53] D. A. Singleton, B. E. Schulmeier, C. Hang, A. A. Thomas, S.-W. Leung, and S. R. Merrigan, Tethraedron **57**, 5149 (2001).

[54] E. Goldstein, B. Beno, and K. N. Houk, J. Am. Chem. Soc. **118**, 6036 (1996).

[55] S. Sakai, J. Phys. Chem. A **104**, 922 (2000).

[56] L. R. Domingo and J. A. Saez, Org. Biomol. Chem. **7**, 3576 (2009).

[57] P. J. Donoghue and O. Wiest, Chem. Eur. J. **12**, 7018 (2006).

[58] V. I. Minkin, Pure Appl. Chem. **71**, 1919 (1999).

[59] K. Black, P. Liu, L. Xu, C. Doubleday, and K. N. Houk, PNAS **109**, 12860 (2012).

[60] R. Hoffmann and R. B. Woodward, Acc. Chem. Res. **1**, 17 (1968).

[61] S. Wilsey, K. N. Houk, and A. H. Zewail, J. Am. Chem. Soc. **121**, 5772 (1999).

[62] G. Gayatri, Indian J. Chem. **50A**, 1579 (2011).

[63] M. Linder and T. Brinck, Phys. Chem. Chem. Phys. **15**, 5108 (2013).

[64] O. Wiest, D. C. Montiel, and K. N. Houk, J. Phys. Chem. A **101**, 8378 (1997).

[65] V. Guner, K. S. Khuong, A. G. Leach, P. S. Lee, M. D. Bartberger, and K. N. Houk, J. Phys. Chem. A **107**, 11445 (2003).

[66] P. D. Mezei, G. I. Csonka, and M. Kállay, J. Chem. Theory Comput. **11**, 2879 (2015).

[67] H. Lischka, E. Ventura, and M. Dallos, Chem. Phys. Chem. **5**, 1365 (2004).

[68] D. H. Ess and K. N. Houk, J. Phys. Chem. A **109**, 9542 (2005).

[69] S.-Y. Tang, J. Shi, and Q.-X. Guo, Org. Biomol. Chem. **10**, 2673 (2012).

[70] D. J. Bellville and N. L. Bauld, J. Am. Chem. Soc. **104**, 2665 (1982).

[71] N. L. Bauld, D. J. Bellville, B. Harirchian, K. T. Lorenz, R. A. Pabon, D. W. Reynolds, D. D. Wirth, H.-S. Chiou, and B. K. Marsh, Acc. Chem. Res. **20**, 371 (1987).

[72] O. Wiest, E. Steckhan, and F. Grein, J. Org. Chem. **57**, 4034 (1992).

[73] N. L. Bauld, D. J. Bellville, R. Pabon, R. Chelsky, and G. Green, J. Am. Chem. Soc. **105**, 2378 (1983).

[74] K. Chockalingam, M. Pinto, and N. L. Bauld, J. Am. Chem. Soc. **112**, 447 (1990).

[75] G. Bouchoux, J. Salpin, and F. Turecek, Rapid Commun. Mass Spectrom. **8**, 325 (1994).

[76] M. Hofmann and H. F. Shaefer III, J. Am. Chem. Soc. **121**, 6719 (1999).

[77] G. Bouchoux, J.-Y. Salpin, and M. Yáñez, J. Phys. Chem. A **108**, 9853 (2004).

[78] J. S. J. Tan, V. Hirvonen, and R. S. Paton, Org. Lett. **20**, 2821 (2018).

[79] Z. Wang, J. S. Hirschi, and D. A. Singleton, Angew. Chem. Int. Ed. **48**, 9156 (2009).

[80] F. Liu, Z. Yang, Y. Mei, and K. N. Houk, J. Phys. Chem. B **120**, 6250 (2016).

[81] J. Soto-Delgado, R. A. Tapia, and J. Torras, J. Chem. Theory Comput. **12**, 4735 (2016).

[82] S. Brickel and M. Meuwly, J. Phys. Chem. A **121**, 5079 (2017).

[83] D. Koner, R. J. Bemish, and M. Meuwly, J. Chem. Phys. **149**, 094305 (2018).

[84] O. Denis-Alpizar, R. J. Bemish, and M. Meuwly, Phys. Chem. Chem. Phys. **19**, 2392 (2017).

[85] S. Willitsch, Adv. Chem. Phys. **162**, 307 (2017).

[86] F. Jensen, *Introduction to Computational Chemistry*, 2 ed. (John Wiley & Sons, Inc., USA, 2006).

[87] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover, Mineola, 1996).

[88] B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).

[89] E. G. Lewars, *Computational Chemistry. Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, 2 ed. (Springer, New York, 2011).

[90] R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. **54**, 724 (1971).

[91] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[92] J. J. P. Stewart, J. Mol. Model. **13**, 1173 (2007).

[93] J. J. P. Stewart, J. Mol. Model. **19**, 1 (2013).

[94] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2 ed. (Wiley-VCH Verlag GmbH, 2001).

[95] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2 ed. (Wiley, England, 2004).

[96] J. Harris, Phys. Rev. A **29**, 1648 (1984).

[97] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[98] Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).

[99] T. Nagy, J. Y. Reyes, and M. Meuwly, J. Chem. Theory Comput. **10**, 1366 (2014).

[100] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. WiÃşrkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).

[101] H. C. Urey and C. A. Vradley Jr, Phys. Rev. **38**, 1969 (1931).

[102] B. Huang and O. A. von Lilienfeld, arXiv preprint arXiv:1707.04146 (2017).

[103] O. T. Unke and M. Meuwly, J. Chem. Theory Comput. **15**, 3678 (2019).

[104] H. P. Hratchian and H. B. Schlegel, J. Chem. Theory Comput. **1**, 61 (2005).

[105] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus., J. Comput. Chem. **4**, 187 (1983).

[106] S. Fischer and M. Karplus, Chem. Phys. Lett. **194**, 252 (1992).

[107] S. Maeda, K. Ohno, and K. Morokuma, Phys. Chem. Chem. Phys. **15**, 3683 (2013).

[108] K. Ohno and S. Maeda, Chem. Phys. Lett. **384**, 277 (2004).

[109] R. D. Levine, *Molecular Reaction Dynamics* (Cambridge University Press, Cambridge, 2005).

[110] W. H. Miller, J. Am. Chem. Soc. **101**, 6810 (1979).

[111] A. Fernandez-Ramos, B. A. Ellingson, B. C. Garrett, and D. G. Truhlar, *Variational Transition State Theory with Multidimensional Tunneling. In Reviews in Computational Chemistry* (John Wiley & Sons, Inc, 2007).

[112] R. G. Gilbert and S. C. Smith, *Theory of unimolecular and recombination reactions* (Blackwell Scientific Publications, Oxford, 1990).

[113] S. E. Stein and B. S. Rabinovitch, J Chem. Phys. **58**, 2438 (1973).

[114] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

[115] R. W. Hockney, S. P. Goel, and J. Eastwood, J. Comp. Phys. **14**, 148 (1974).

[116] R. B. Bernstein, *Atom-Molecule Collision Theory. A Guide for the Experimentalists* (Plenum Press, Columbia University, New York, 1979).

[117] J. Nelder and R. Mead, Chem. Phys. **7**, 308 (1965).

[118] J. Lee, H. A. Scheraga, and S. Rackovsky, J. Comput. Chem. **18**, 1222 (1997).

[119] O. T. Unke, M. Devereux, and M. Meuwly, J. Chem. Phys. **147**, 161712 (2017).

[120] M. N. Eberlin and R. G. Cooks, J. Am. Chem. Soc. **115**, 9226 (1993).

[121] M. J. Frisch *et al.*, Gaussian09 Revision E.01, Gaussian Inc. Wallingford CT 2009.

[122] S. Maeda, Y. Harabuchi, Y. Osada, T. Taketsugu, K. Morokuma, and K. Ohno, GRRM14, 2013.

[123] F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. D. Vico, I. F. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. L. Manni, H. Lischka, D. Ma, P. A. Malmqvist, T. Muller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata, and R. Lindh, J. Comput. Chem. **37**, 506 (2016).

[124] J. R. Barker, T. L. Nguyen, J. F. Stanton, C. Aieta, M. Ceotto, F. Gabas, T. J. D. Kumar, C. G. L. Li, L. L. Lohr, A. Maranzana, N. F. Ortiz, J. M. Preses, and P. J. Stimac, MultiWell-2016 Software Suite, University of Michigan, Ann Arbor, Michigan, USA, 2016.

[125] Maple 2015.0, Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario.

[126] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II*, 2 ed. (Springer, New York, 1996).

[127] K. Alder and G. Stein, Justus Liebigs Ann. Chem. **525**, 221 (1936).

[128] T. J. Lee and P. R. Taylor, Int. J. Quantum Chem. **36**, 199 (1989).

[129] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus., J. Comput. Chem. **30**, 1545 (2009).

[130] V. Zoete, M. A. Cuendet, A. Grosdidier, and O. Michielin, J. Chem. Phys. **32**, 2359 (2011).

[131] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, J. Phys.: Condens. Matter **29**, 273002 (2017).

[132] J. J. P. Stewart, Mopac2016, 2016, `http://openmopac.net/`, Stewart Computational Chemistry, Colorado Springs, CO, USA.

[133] J. Behler, J. Phys. Condens. Matter **26**, 183001 (2014).

[134] P. W. Atkins, *Physical Chemistry*, Fifth ed. (Oxford University Press, 1994).

[135] U. Rivero, M. Meuwly, and S. Willitsch, Chem. Phys. Lett. **683**, 598 (2017).

[136] J. Behler, Int. J. Quantum Chem. **115**, 1032 (2015).

[137] U. Rivero, O. T. Unke, M. Meuwly, and S. Willitsch, 2019, https://doi.org/10.5281/zenodo.3291503.

[138] I. Fernández and F. M. Bickelhaupt, J. Comput. Chem. **35**, 371 (2014).

[139] O. T. Unke, S. Brickel, and M. Meuwly, J. Chem. Phys. **150**, 074107 (2019).

[140] H. Guo and B. Jiang, Acc. Chem. Res. **47**, 3679 (2014).

[141] J. C. Polanyi, Acc. Chem. Res. **5**, 161 (1972).

[142] Z. Zhang, Y. Zhou, D. H. Zhang, G. Czakó, and J. M. Bowman, J. Phys. Chem. Lett. **3**, 3416 (2012).

[143] B. Jiang, J. Li, and H. Guo, J. Chem. Phys. **140**, 034112 (2014).

[144] A. Li and H. Guo, J. Phys. Chem. A **118**, 11168 (2014).

[145] H. Song, A. Li, H. Guo, Y. Xu, B. Xiong, Y.-C. Chang, and C. Y. Ng, Phys. Chem. Chem. Phys. **18**, 22509 (2016).

[146] A. Li and H. Guo, J. Chem. Phys. **140**, 224313 (2014).

[147] A. J. Orr-Ewing, Chem. Soc. Rev. **46**, 7597 (2017).

[148] S. Brickel and M. Meuwly, J. Phys. Chem. B **123**, 448 (2019).

[149] D. Rösch, S. Willitsch, Y.-P. Chang, and J. Küpper, J. Chem. Phys. **140**, 124202 (2014).

[150] D. A. Horke, Y.-P. Chang, K. Dlugolecki, and J. Küpper, Angew. Chem. Int. Ed. **53**, 11965 (2014).

[151] Y.-P. Chang, D. A. Horke, S. Trippel, and J. Küpper, Int. Rev. Phys. Chem. **34**, 557 (2015).

[152] D. Rösch, H. Gao, A. Kilaj, and S. Willitsch, EPJ Tech. Instrum. **3**, 5 (2016).

[153] D. Clary, J. Chem. Soc., Faraday Trans. 2 **83**, 139 (1987).

[154] T. Stoecklin, D. C. Clary, and A. Palma, J. Chem. Soc. Faraday Trans. **88**, 901 (1992).

Appendix A

MS-ARMD model for the Diels-Alder reaction DBB + MA

		Reactant FF		Product FF	
Atom 1 #	Atom 2 #	$k/2$ [kcal/mol/Å^2]	r_e [Å]	$k/2$ [kcal/mol/Å^2]	r_e [Å]
2	4	1093.65	1.19577	1310.44	1.19499
3	5	1093.65	1.19577	1310.44	1.19499
6	8	411.539	1.08513	464.889	1.09667
7	9	411.539	1.08513	464.889	1.09667
11	15	403.908	1.08680	464.889	1.09667
12	17	403.908	1.08680	464.889	1.09667
11	13	403.908	1.08680	464.889	1.09667
10	12	403.908	1.08680	464.889	1.09667

Table A.1: Harmonic bond parameters of the MS-ARMD force fields.

		Reactant FF			Product FF		
Atom 1 #	Atom 2 #	D_e [kcal/mol]	r_e [Å]	β [1/Å]	D_e [kcal/mol]	r_e [Å]	β [1/Å]
1	2	85.1832	1.39166	1.99423	62.7150	1.39718	1.98961
1	3	85.1832	1.39166	1.99423	62.7150	1.39718	1.98961
2	6	164.953	1.50368	1.40025	72.0637	1.52370	1.98423
3	7	164.953	1.50368	1.40025	72.0637	1.52370	1.98423
6	7	194.950	1.33603	1.99291	183.946	1.50919	1.07343
14	16	205.845	1.45846	1.40818	360.962	1.34254	1.48736
11	14	353.111	1.32505	1.47075	95.4170	1.50611	1.91729
12	16	353.111	1.32505	1.47075	95.4170	1.50611	1.91729
7	12	X	X	X	183.946	1.50919	1.07343
6	11	X	X	X	183.946	1.50919	1.07343
14	18	65.3971	1.89921	1.87225	68.8197	1.88376	1.96818
16	19	65.3971	1.89921	1.87225	68.8197	1.88376	1.96818

Table A.2: Morse bond parameters of the MS-ARMD force fields. "X" indicates that this parameter is not needed.

			Reactant FF		Product FF	
Atom 1 #	Atom 2 #	Atom 3 #	$k/2$ [kcal/mol/radian^2]	θ_e [degree]	$k/2$ [kcal/mol/radian^2]	θ_e [degree]
2	1	3	64.5319	105.978	153.226	113.237
1	2	4	99.6847	128.737	109.650	131.035
1	2	6	135.436	109.951	99.6751	119.974
4	2	6	77.1982	137.451	78.9705	140.440
1	3	5	99.6847	128.737	109.650	131.035
1	3	7	135.436	109.951	99.6751	119.974
5	3	7	77.1982	137.451	78.9705	140.440
2	6	7	95.5494	108.805	64.0062	103.976
2	6	8	38.5239	124.711	50.4823	105.428
2	6	11	X	X	64.0062	103.976
7	6	8	20.1303	129.751	56.7929	108.108
7	6	11	X	X	77.4704	100.923
8	6	11	X	X	56.7929	108.108
3	7	6	95.5494	108.805	64.0062	103.976
3	7	9	38.5239	124.711	50.4823	105.428
3	7	12	X	X	64.0062	103.976
6	7	9	20.1303	129.751	56.7929	108.108
6	7	12	X	X	77.4704	100.923
9	7	12	X	X	56.7929	108.108
6	11	13	X	X	56.7929	108.108
6	11	14	X	X	43.0251	101.421
6	11	15	X	X	56.7929	108.108
13	11	14	42.9328	123.393	55.5841	110.000
13	11	15	26.6835	123.557	47.2662	107.155
14	11	15	42.9328	123.393	55.5841	110.000
7	12	10	X	X	56.7929	108.108
7	12	16	X	X	43.0251	101.421
7	12	17	X	X	56.7929	108.108
10	12	16	42.9328	123.393	55.5841	110.000
10	12	17	26.6835	123.557	47.2662	107.155
16	12	17	42.9328	123.393	55.5841	110.000
11	14	16	38.6304	128.237	55.9672	122.309
11	14	18	64.6768	127.307	69.3366	115.166
16	14	18	55.6080	121.871	87.6601	124.293
12	16	14	38.6304	128.237	55.9672	122.309
12	16	19	64.6768	127.307	69.3366	115.166
14	16	19	55.6080	121.871	87.6601	124.293

Table A.3: Angle parameters of the MS-ARMD force fields. "X" indicates that this parameter is not needed.

					Reactant FF	Product FF	
Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol]	k [kcal/mol]	ϕ [degree]
1	2	6	7	1	-0.071	-0.456229	0.00
1	2	6	7	2	4.79788	1.33681	180.00
1	2	6	7	3	X	0.725632	0.00
1	2	6	8	1	0.179	X	0.00
1	2	6	8	2	0.249864	1.52218	180.00
1	2	6	8	3	0.097	1.52546	0.00
1	2	6	11	1	X	-0.456229	0.00
1	2	6	11	2	X	1.33681	180.00
1	2	6	11	3	X	0.725632	0.00
1	3	7	6	1	-0.071	-0.456229	0.00
1	3	7	6	2	4.79788	1.33681	180.00
1	3	7	6	3	X	0.725632	0.00
1	3	7	9	1	0.179	X	0.00
1	3	7	9	2	0.249864	1.52218	180.00
1	3	7	9	3	0.097	1.52546	0.00
1	3	7	12	1	X	-0.456229	0.00
1	3	7	12	2	X	1.33681	180.00
1	3	7	12	3	X	0.725632	0.00
2	1	3	5	1	0.701169	2.91529	0.00
2	1	3	5	2	6.21522	1.47165	180.00
2	1	3	5	3	0.298850	0.909140	0.00
2	1	3	7	2	7.17573	4.31208	180.00
2	6	7	3	1	X	-1.41346	0.00
2	6	7	3	2	8.19928	X	180.00
2	6	7	3	3	X	-2.99571	0.00
2	6	7	9	1	X	1.82442	0.00
2	6	7	9	2	6.34212	-2.00004	180.00
2	6	7	12	1	X	-1.95087	0.00
2	6	7	12	2	X	-0.959448	180.00
2	6	7	12	3	X	0.307973	0.00
2	6	11	13	1	X	1.82442	0.00
2	6	11	13	2	X	-2.00004	180.00
2	6	11	14	3	X	-2.07692	0.00
2	6	11	15	1	X	1.82442	0.00
2	6	11	15	2	X	-2.00004	180.00
3	1	2	4	1	0.701169	2.91529	0.00
3	1	2	4	2	6.21522	1.47165	180.00
3	1	2	4	3	0.298850	0.909140	0.00
3	1	2	6	2	7.17573	0.909140	180.00
3	7	6	8	1	X	1.82442	0.00
3	7	6	8	2	6.34212	-2.00004	180.00
3	7	6	11	1	X	-1.95087	0.00
3	7	6	11	2	X	-0.959448	180.00
3	7	6	11	3	X	0.307973	0.00
3	7	12	10	1	X	1.82442	0.00
3	7	12	10	2	X	-2.00004	180.00
3	7	12	16	3	X	-2.07692	0.00
3	7	12	17	1	X	1.82442	0.00
3	7	12	17	2	X	-2.00004	180.00
4	2	6	7	1	0.181	-0.224498E-01	0.00
4	2	6	7	2	-0.999624	1.23963	180.00
4	2	6	7	3	X	-0.838198E-01	0.00
4	2	6	8	1	X	-2.02295	0.00
4	2	6	8	2	2.45356	-0.342985	180.00
4	2	6	8	3	X	-0.853726E-01	0.00

					Reactant FF	Product FF	
Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol]	k [kcal/mol]	ϕ [degree]
4	2	6	11	1	X	-0.224498E-01	0.00
4	2	6	11	2	X	1.23963	180.00
4	2	6	11	3	X	-0.838198E-01	0.00
5	3	7	6	1	0.181	-0.224498E-01	0.00
5	3	7	6	2	-0.999624	1.23963	180.00
5	3	7	6	3	X	-0.838198E-01	0.00
5	3	7	9	1	X	-2.02295	0.00
5	3	7	9	2	2.45356	-0.342985	180.00
5	3	7	9	3	X	-0.853726E-01	0.00
5	3	7	12	1	X	-0.224498E-01	0.00
5	3	7	12	2	X	1.23963	180.00
5	3	7	12	3	X	-0.838198E-01	0.00
6	7	12	10	1	X	-1.73754	0.00
6	7	12	10	2	X	-0.739873	180.00
6	7	12	10	3	X	0.842630E-01	0.00
6	7	12	16	1	X	-0.800732	0.00
6	7	12	16	2	X	-1.97784	180.00
6	7	12	16	3	X	0.824677	0.00
6	7	12	17	1	X	-1.73754	0.00
6	7	12	17	2	X	-0.739873	180.00
6	7	12	17	3	X	0.842630E-01	0.00
6	11	14	16	1	X	-0.758787	0.00
6	11	14	16	2	X	-0.416878	180.00
6	11	14	16	3	X	-0.739873	0.00
6	11	14	18	1	X	0.000	0.00
7	6	11	13	1	X	-1.73754	0.00
7	6	11	13	2	X	-0.739873	180.00
7	6	11	13	3	X	0.842630E-01	0.00
7	6	11	14	1	X	-0.800732	0.00
7	6	11	14	2	X	-1.97784	180.00
7	6	11	14	3	X	0.824677	0.00
7	6	11	15	1	X	-1.73754	0.00
7	6	11	15	2	X	-0.739873	180.00
7	6	11	15	3	X	0.842630E-01	0.00
7	12	16	14	1	X	-0.800732	0.00
7	12	16	14	2	X	-0.416878	180.00
7	12	16	14	3	X	-0.739873	0.00
7	12	16	19	1	X	0.000	0.00
8	6	7	9	1	X	0.343877	0.00
8	6	7	9	2	2.34735	-1.99591	180.00
8	6	7	9	3	X	-0.136668	0.00
8	6	7	12	1	X	-1.73754	0.00
8	6	7	12	2	X	-0.739873	180.00
8	6	7	12	3	X	0.842630E-01	0.00
8	6	11	13	1	X	0.343877	0.00
8	6	11	13	2	X	-1.99591	180.00
8	6	11	13	3	X	-0.136668	0.00
8	6	11	14	1	X	1.47393	0.00
8	6	11	14	2	X	-0.225128	180.00
8	6	11	14	3	X	0.355477	0.00
8	6	11	15	1	X	0.343877	0.00
8	6	11	15	2	X	-1.99591	180.00
8	6	11	15	3	X	-0.136668	0.00

					Reactant FF	Product FF	
Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol]	k [kcal/mol]	ϕ [degree]
9	7	6	11	1	X	-1.73754	0.00
9	7	6	11	2	X	-0.739873	180.00
9	7	6	11	3	X	0.842630E-01	0.00
9	7	12	10	1	X	0.343877	0.00
9	7	12	10	2	X	-1.99591	180.00
9	7	12	10	3	X	-0.136668	0.00
9	7	12	16	1	X	1.47393	0.00
9	7	12	16	2	X	-0.225128	180.00
9	7	12	16	3	X	0.355477	0.00
9	7	12	17	1	X	0.343877	0.00
9	7	12	17	2	X	-1.99591	180.00
9	7	12	17	3	X	-0.136668	0.00
10	12	16	14	1	X	-0.775003E-01	0.00
10	12	16	14	2	4.77285	-0.225128	180.00
10	12	16	14	3	X	-0.876922	0.00
10	12	16	19	1	X	0.000	0.00
10	12	16	19	2	6.16283	X	180.00
11	6	7	12	1	X	-0.314136	0.00
11	6	7	12	2	X	3.05348	180.00
11	6	7	12	3	X	-0.796867	0.00
11	14	16	12	1	-5.25896	-1.28302	0.00
11	14	16	12	2	-0.525596	6.94818	180.00
11	14	16	12	3	0.085522	X	0.00
11	14	16	19	2	0.143296	6.94818	180.00
12	16	14	18	2	0.143296	6.94818	180.00
13	11	14	16	1	X	-0.775003E-01	0.00
13	11	14	16	2	4.77285	-0.225128	180.00
13	11	14	16	3	X	-0.876922	0.00
13	11	14	18	1	X	0.000	0.00
13	11	14	18	2	6.16283	X	180.00
14	16	12	17	1	X	-0.775003E-01	0.00
14	16	12	17	2	4.26042	-0.225128	180.00
14	16	12	17	3	X	-0.876922	0.00
15	11	14	16	1	X	-0.775003E-01	0.00
15	11	14	16	2	4.77285	-0.225128	180.00
15	11	14	16	3	X	-0.876922	0.00
15	11	14	18	1	X	0.000	0.00
15	11	14	18	2	6.16283	X	180.00
17	12	16	19	1	X	0.000	0.00
17	12	16	19	2	6.16283	X	180.00
18	14	16	19	2	0.350057	6.94818	180.00

Table A.4: Dihedral parameters of the MS-ARMD force fields. "X" indicates that this parameter is not needed.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$[Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$[Å]
1	-0.300000	0.725240	1.52534	X	X
2	0.705600	0.100119E-05	3.49075	X	X
3	0.705600	0.100119E-05	3.49075	X	X
4	-0.570000	0.116541E-01	1.81590	0.12	1.40
5	-0.570000	0.116541E-01	1.81590	0.12	1.40
6	-0.135600	0.180888E-01	2.02067	X	X
7	-0.135600	0.180888E-01	2.02067	X	X
8	0.150000	0.146019	1.26142	X	X
9	0.150000	0.146019	1.26142	X	X
10	0.150000	0.146019	1.26142	X	X
11	-0.300000	0.180888E-01	2.02067	X	X
12	-0.300000	0.180888E-01	2.02067	X	X
13	0.150000	0.146019	1.26142	X	X
14	0.110000	0.180888E-01	2.02067	X	X
15	0.150000	0.146019	1.26142	X	X
16	0.110000	0.180888E-01	2.02067	X	X
17	0.150000	0.146019	1.26142	X	X
18	-0.110000	6.09988	1.75471	X	X
19	-0.110000	6.09988	1.75471	X	X
Atom 1 #	Atom 2#	ϵ_i [kcal/mol]	$R_{min}/2$[Å]	n	m
18	4	2.03805	3.29190	13.1468	16.0708
18	5	2.03805	3.29190	13.1468	16.0708
19	4	2.03805	3.29190	13.1468	16.0708
19	5	2.03805	3.29190	13.1468	16.0708
11	6	1.84792	1.87951	3.24440	5.14600
12	7	1.84792	1.87951	3.24440	5.14600
11	7	1.84792	1.87951	3.24440	5.14600
12	6	1.84792	1.87951	3.24440	5.14600

Table A.5: Non bonded parameters of the MS-ARMD reactant force field. "X" indicates that this parameter is not needed.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$[Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$[Å]
1	-0.300000	0.152100	1.770000	X	X
2	0.659000	0.110000	2.000000	X	X
3	0.659000	0.110000	2.000000	X	X
4	-0.570000	0.120000	1.700000	0.12	1.40
5	-0.570000	0.120000	1.700000	0.12	1.40
6	0.610000E-01	0.055000	2.175000	0.01	1.90
7	0.610000E-01	0.055000	2.175000	0.01	1.90
8	0.00000	0.022000	1.320000	X	X
9	0.00000	0.022000	1.320000	X	X
10	0.00000	0.022000	1.320000	X	X
11	0.138200	0.055000	2.175000	0.01	1.90
12	0.138200	0.055000	2.175000	0.01	1.90
13	0.00000	0.022000	1.320000	X	X
14	-0.282000E-01	0.068000	2.090000	X	X
15	0.00000	0.022000	1.320000	X	X
16	-0.282000E-01	0.068000	2.090000	X	X
17	0.00000	0.022000	1.320000	X	X
18	-0.110000	5.48335	1.54929	X	X
19	-0.110000	5.48335	1.54929	X	X

Table A.6: Non bonded parameters of the MS-ARMD products force field. "X" indicates that this parameter is not needed.

The barrier region connecting the reactant and product state is described by two GAPOs

$$\Delta V^{ij}_{GAPO,k}(x) = exp\left(-\frac{(\Delta V_{ij}(x) - V^0_{ij,k})^2}{2\sigma^2_{ij,k}}\right) \times \sum_{l=0}^{m_{ij,k}} a_{ij,kl}(\Delta V_{ij}(x) - V^0_{ij,k})^l$$

with the parameters summarized in Table A.7.

k	$V^0_{ij,k}$	$\sigma_{ij,k}$	$a_{ij,k0}$	$a_{ij,k1}$
1	-3.2889475989E+01	7.7839312146E+01	-4.0222992796E+01	-2.4658081680E-01
0	1.4151964786E+00	2.6626692901E+01	-9.5075322684E+00	

Table A.7: GAPO parameters: i labels the reactant, j labels the product, $V^0_{ij,k}$ is the center of the Gaussian function (in kcal/mol), and $\sigma_{ij,k}$ the width of the Gaussian (in kcal/mol). a_{ij} is the polynomial coefficient in kcal/mol $^{1-j}$, j = 0, 1.

Appendix B

MS-ARMD model for the Diels-Alder reaction $DBB^+ + MA$

Atoms		Reactant FF		Intermediate FF		Product FF	
1 #	2 #	$k/2$	r_e	$k/2$	r_e	$k/2$	r_e
2	4	1093.65	1.19577	1398.03	1.19312	1029.89	1.19351
3	5	1093.65	1.19577	1398.03	1.19312	1029.89	1.19351
6	8	411.539	1.08513	650.507	1.08709	411.200	1.09308
7	9	411.539	1.08513	650.507	1.08709	411.200	1.09308
11	15	391.619	1.08580	650.507	1.08709	411.200	1.09308
12	17	391.619	1.08580	650.507	1.08709	411.200	1.09308
11	13	391.619	1.08580	650.507	1.08709	411.200	1.09308
10	12	391.619	1.08580	650.507	1.08709	411.200	1.09308

Table B.1: Harmonic bond parameters of the MS-ARMD force fields. $k/2$ has units of kcal/mol/Å^2 and r_e is in Å.

Atoms		Reactant FF			Intermediate FF			Product FF		
1 #	2 #	D_e	r_e	β	D_e	r_e	β	D_e	r_e	β
14	18	187.087	1.83777	1.10111	85.9534	1.81583	1.20270	91.2590	1.80871	1.89304
16	19	187.087	1.83777	1.10111	85.9534	1.81583	1.20270	91.2590	1.80871	1.89304
1	2	85.1832	1.39166	1.99423	101.912	1.34206	1.87376	474.144	1.37373	1.00021
1	3	85.1832	1.39166	1.99423	101.912	1.34206	1.87376	474.144	1.37373	1.00021
2	6	164.953	1.50368	1.40025	63.8103	1.55906	1.18897	189.327	1.52650	1.38405
3	7	164.953	1.50368	1.40025	63.8103	1.55906	1.18897	189.327	1.52650	1.38405
14	16	564.329	1.40983	0.896631	327.948	1.41861	1.44211	310.137	1.42532	1.12287
11	14	189.781	1.36281	1.63991	76.1220	1.32884	1.01309	236.306	1.49451	1.24374
12	16	189.781	1.36281	1.63991	76.1220	1.32884	1.01309	236.306	1.49451	1.24374
6	7	194.950	1.33603	1.99291	154.586	1.50359	1.94850	58.3539	1.47745	1.99994
7	12	X	X	X	116.127	1.51263	1.01790	58.3539	1.47745	1.99994
6	11	X	X	X	X	X	X	58.3539	1.47745	1.99994

Table B.2: Morse bond parameters of the MS-ARMD force fields. "X" indicates that this parameter is not needed. D_e is in kcal/mol, r_e in Å and β in $Å^{-1}$.

Atoms			Reactant FF		Intermediate FF		Product FF	
1 #	2 #	3 #	$k/2$	θ_e	$k/2$	θ_e	$k/2$	θ_e
2	1	3	64.5319	105.978	135.149	97.7709	109.386	111.098
1	2	4	99.6847	128.737	165.070	130.560	114.817	135.968
1	2	6	135.436	109.951	277.466	107.951	134.612	120.559
4	2	6	77.1982	137.451	78.2342	143.084	71.8608	148.445
1	3	5	99.6847	128.737	165.070	130.560	114.817	135.968
1	3	7	135.436	109.951	277.466	107.951	134.612	120.559
5	3	7	77.1982	137.451	78.2342	143.084	71.8608	148.445
2	6	7	95.5494	108.805	134.514	110.650	89.4152	108.264
2	6	8	38.5239	124.711	67.2617	121.288	61.0991	109.775
2	6	11	X	X	X	X	89.4152	108.264
7	6	8	20.1303	129.751	64.9256	125.690	54.6427	108.583
7	6	11	X	X	X	X	82.9790	98.3030
8	6	11	X	X	X	X	54.6427	108.583
3	7	6	95.5494	108.805	134.514	110.650	89.4152	108.264
3	7	9	38.5239	124.711	67.2617	121.288	61.0991	109.775
3	7	12	X	X	134.514	110.650	89.4152	108.264
6	7	9	20.1303	129.751	64.9256	125.690	54.6427	108.583
6	7	12	X	X	35.4515	133.262	82.9790	98.3030
9	7	12	X	X	64.9256	125.690	54.6427	108.583
6	11	13	X	X	X	X	54.6427	108.583
6	11	14	X	X	X	X	65.2956	104.164
6	11	15	X	X	X	X	54.6427	108.583
13	11	14	44.8358	132.443	123.161	121.235	58.4988	108.370
13	11	15	30.0806	138.270	106.624	119.564	44.9477	107.173
14	11	15	44.8358	132.443	123.161	121.235	58.4988	108.370
7	12	10	X	X	64.9256	125.690	54.6427	108.583
7	12	16	X	X	46.2212	134.655	65.2956	104.164
7	12	17	X	X	64.9256	125.690	54.6427	108.583
10	12	16	44.8358	132.443	123.161	121.235	58.4988	108.370
10	12	17	30.0806	138.270	106.624	119.564	44.9477	107.173
16	12	17	44.8358	132.443	123.161	121.235	58.4988	108.370
11	14	16	112.001	124.423	57.8435	132.115	92.0564	132.027
11	14	18	43.8118	136.630	46.9468	138.569	51.8620	130.256
16	14	18	50.3729	127.032	22.9482	148.260	81.7961	128.923
12	16	14	112.001	124.423	57.8435	132.115	92.0564	132.027
12	16	19	43.8118	136.630	46.9468	138.569	51.8620	130.256
14	16	19	50.3729	127.032	22.9482	148.260	81.7961	128.923

Table B.3: Angle parameters of the MS-ARMD force fields. "X" indicates that this parameter is not needed. $k/2$ is in kcal/mol/radian^2, θ_e in degree

Atoms					Reactant FF	Intermediate FF	Product FF	
1 #	2 #	3 #	4 #	N	k	k	k	ϕ
1	2	6	7	1	X	-4.14146	-0.523798	0.00
1	2	6	7	2	4.79788	8.55575	1.72536	180.00
1	2	6	7	3	X	-1.81563	0.267926	0.00
1	2	6	8	2	0.249864	2.35984	2.15011	180.00
1	2	6	8	3	X	0.221796	1.66411	0.00
1	2	6	11	1	X	X	-0.523798	0.00
1	2	6	11	2	X	X	1.72536	180.00
1	2	6	11	3	X	X	0.267926	0.00
1	3	7	6	1	X	-4.14146	-0.523798	0.00
1	3	7	6	2	4.79788	8.55575	1.72536	180.00
1	3	7	6	3	X	-1.81563	0.267926	0.00
1	3	7	9	2	0.249864	2.35984	2.15011	180.00
1	3	7	9	3	X	0.221796	1.66411	0.00
1	3	7	12	1	X	-4.14146	-0.523798	0.00
1	3	7	12	2	X	8.55575	1.72536	180.00
1	3	7	12	3	X	-1.81563	0.267926	0.00
2	1	3	5	1	0.701169	-5.58519	-0.126467	0.00
2	1	3	5	2	6.21522	-9.49140	-5.49755	180.00
2	1	3	5	3	0.298850	12.7562	3.34467	0.00
2	1	3	7	2	7.17573	-1.12647	4.91803	180.00
2	6	7	3	1	X	-3.95005	-1.29628	0.00
2	6	7	3	2	8.19928	-1.12703	X	X
2	6	7	3	3	X	0.390811	0.330764	0.00
2	6	7	9	1	X	1.45039	0.336052	0.00
2	6	7	9	2	6.34212	-0.211970	-2.95419	180.00
2	6	7	12	1	X	-5.40684	-3.89326	0.00
2	6	7	12	2	X	0.454822	-2.88140	180.00
2	6	7	12	3	X	-0.0439145	-0.444273	0.00
2	6	11	13	1	X	X	0.336052	0.00
2	6	11	13	2	X	X	-2.95419	180.00
2	6	11	14	3	X	X	0.035506	0.00
2	6	11	15	1	X	X	0.336052	0.00
2	6	11	15	2	X	X	-2.95419	180.00
3	1	2	4	1	0.701169	-5.58519	-0.126467	0.00
3	1	2	4	2	6.21522	-9.49140	-5.49755	180.00
3	1	2	4	3	0.298850	12.7562	3.34467	0.00
3	1	2	6	2	7.17573	12.7562	3.34467	180.00
3	7	6	8	1	X	1.45039	0.336052	0.00
3	7	6	8	2	6.34212	-0.211970	-2.95419	180.00
3	7	6	11	1	X	X	-3.89326	0.00
3	7	6	11	2	X	X	-2.88140	180.00
3	7	6	11	3	X	X	-0.444273	0.00
3	7	12	10	1	X	1.45039	0.336052	0.00
3	7	12	10	2	X	-0.211970	-2.95419	180.00
3	7	12	16	3	X	1.51659	0.035506	0.00
3	7	12	17	1	X	1.45039	0.336052	0.00
3	7	12	17	2	X	-0.211970	-2.95419	180.00
4	2	6	7	1	X	-9.15296	-2.89358	0.00
4	2	6	7	2	-0.999624	-1.25200	1.75806	180.00

Atoms					Reactant FF	Intermediate FF	Product FF	
1 #	2 #	3 #	4 #	N	k	k	k	ϕ
4	2	6	8	1	X	-10.0609	-2.16831	0.00
4	2	6	8	2	2.45356	-2.58269	1.09179	180.00
4	2	6	8	3	X	-1.46073	0.192876	0.00
4	2	6	11	1	X	X	-2.89358	0.00
4	2	6	11	2	X	X	1.75806	180.00
4	2	6	11	3	X	X	0.230563	0.00
5	3	7	6	1	X	-9.15296	-2.89358	0.00
5	3	7	6	2	-0.999624	-1.25200	1.75806	180.00
5	3	7	6	3	X	-0.997410	0.230563	0.00
5	3	7	9	1	X	-10.0609	-2.16831	0.00
5	3	7	9	2	2.45356	-2.58269	1.09179	180.00
5	3	7	9	3	X	-1.46073	0.192876	0.00
5	3	7	12	1	X	-9.15296	-2.89358	0.00
5	3	7	12	2	X	-1.25200	1.75806	180.00
5	3	7	12	3	X	-0.997410	0.230563	0.00
6	7	12	10	1	X	-6.25427	-1.53021	0.00
6	7	12	10	2	X	-0.819813	-0.713956	180.00
6	7	12	10	3	X	-1.40487	-0.0941236	0.00
6	7	12	16	1	X	0.0503902	-1.74289	0.00
6	7	12	16	2	X	1.43493	-2.69254	180.00
6	7	12	16	3	X	5.08310	-2.56013	0.00
6	7	12	17	1	X	-6.25427	-1.53021	0.00
6	7	12	17	2	X	-0.819813	-0.713956	180.00
6	7	12	17	3	X	-1.40487	-0.0941236	0.00
6	11	14	16	1	X	X	-2.46899	0.00
6	11	14	16	2	X	X	0.0047732	180.00
6	11	14	16	3	X	X	-0.713956	0.00
6	11	14	18	1	X	X	0.000	0.00
7	6	11	13	1	X	X	-1.53021	0.00
7	6	11	13	2	X	X	-0.713956	180.00
7	6	11	13	3	X	X	-0.094123	0.00
7	6	11	14	1	X	X	-1.74289	0.00
7	6	11	14	2	X	X	-2.69254	180.00
7	6	11	14	3	X	X	-2.56013	0.00
7	6	11	15	1	X	X	-1.53021	0.00
7	6	11	15	2	X	X	-0.713956	180.00
7	6	11	15	3	X	X	-0.0941236	0.00
7	12	16	14	1	X	1.56618	-1.74289	0.00
7	12	16	14	2	X	-2.45376	0.0047732	180.00
7	12	16	14	3	X	-0.819813	-0.713956	0.00
7	12	16	19	1	X	0.000	0.000	0.00
8	6	7	9	1	X	-3.43664	-0.0510087	0.00
8	6	7	9	2	2.34735	1.24857	-2.28477	180.00
8	6	7	9	3	X	1.78467	0.0546769	0.00
8	6	7	12	1	X	-6.25427	-1.53021	0.00
8	6	7	12	2	X	-0.819813	-0.713956	180.00
8	6	7	12	3	X	-1.40487	-0.094123	0.00
8	6	11	13	1	X	X	-0.051008	0.00
8	6	11	13	2	X	X	-2.28477	180.00

Atoms					Reactant FF	Intermediate FF	Product FF	
1 #	2 #	3 #	4 #	N	k	k	k	ϕ
8	6	11	13	3	X	X	0.0546769	0.00
8	6	11	14	1	X	X	0.861788	0.00
8	6	11	14	2	X	X	0.995909	180.00
8	6	11	14	3	X	X	-1.07686	0.00
8	6	11	15	1	X	X	-0.051008	0.00
8	6	11	15	2	X	X	-2.28477	180.00
8	6	11	15	3	X	X	0.0546769	0.00
9	7	6	11	1	X	X	-1.53021	0.00
9	7	6	11	2	X	X	-0.713956	180.00
9	7	6	11	3	X	X	-0.0941236	0.00
9	7	12	10	1	X	-3.43664	-0.0510087	0.00
9	7	12	10	2	X	1.24857	-2.28477	180.00
9	7	12	10	3	X	1.78467	0.054676	0.00
9	7	12	16	1	X	-1.19475	0.861788	0.00
9	7	12	16	2	X	3.74789	0.995909	180.00
9	7	12	16	3	X	-1.42060	-1.07686	0.00
9	7	12	17	1	X	-3.43664	-0.0510087	0.00
9	7	12	17	2	X	1.24857	-2.28477	180.00
9	7	12	17	3	X	1.78467	0.0546769	0.00
10	12	16	14	1	X	-4.27768	-3.29003	0.00
10	12	16	14	2	3.21391	3.74789	0.995909	180.00
10	12	16	14	3	X	0.960659	-0.693868	0.00
10	12	16	19	1	X	2.13188	0.000	0.00
10	12	16	19	2	2.53826	-3.11705	X	X
11	6	7	12	1	X	X	-0.601273	0.00
11	6	7	12	2	X	X	5.54093	180.00
11	6	7	12	3	X	X	-1.18306	0.00
11	14	16	12	1	-0.818383	-2.27656	-5.28649	0.00
11	14	16	12	2	1.645260	3.43143	2.36666	180.00
11	14	16	12	3	-1.23972	-2.35012	X	X
11	14	16	19	2	3.07771	0.554677	2.36666	180.00
12	16	14	18	2	3.07771	0.554677	2.36666	180.00
13	11	14	16	1	X	-4.27768	-3.29003	0.00
13	11	14	16	2	2.42160	3.74789	0.995909	180.00
13	11	14	16	3	X	0.960659	-0.693868	0.00
13	11	14	18	1	X	0.000	0.000	0.00
13	11	14	18	2	2.53826	-3.11705	X	X
14	16	12	17	1	X	-4.27768	-3.29003	0.00
14	16	12	17	2	3.21391	3.74789	0.995909	180.00
14	16	12	17	3	X	0.960659	-0.693868	0.00
15	11	14	16	1	X	-4.27768	-3.29003	0.00
15	11	14	16	2	2.42160	3.74789	0.995909	180.00
15	11	14	16	3	X	0.960659	-0.693868	0.00
15	11	14	18	1	X	0.000	0.000	0.00
15	11	14	18	2	2.53826	-3.11705	X	X
17	12	16	19	1	X	0.000	0.000	0.00
17	12	16	19	2	2.53826	-3.11705	X	X
18	14	16	19	2	3.13276	3.43143	2.36666	180.00

Table B.4: Dihedral parameters of the MS-ARMD force fields. "X" indicates that this parameter is not needed. k is in kcal/mol and ϕ in degree.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$[Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$[Å]
1	-0.300000	0.203207E-04	2.80541	X	X
2	0.705600	0.667655	0.102478	X	X
3	0.705600	0.667655	0.102478	X	X
4	-0.570000	0.411920E-08	4.77007	0.120	1.40
5	-0.570000	0.411920E-08	4.77007	0.120	1.40
6	-0.135600	0.242868E-01	2.13203	X	X
7	-0.135600	0.242868E-01	2.13203	X	X
8	0.150000	0.218559	0.917005	X	X
9	0.150000	0.218559	0.917005	X	X
10	0.152910	0.218559	0.917005	X	X
11	0.221270	0.242868E-01	2.13203	X	X
12	0.221270	0.242868E-01	2.13203	X	X
13	0.152910	0.218559	0.917005	X	X
14	-0.151030	0.242868E-01	2.13203	X	X
15	0.152910	0.218559	0.917005	X	X
16	-0.151030	0.242868E-01	2.13203	X	X
17	0.152910	0.218559	0.917005	X	X
18	0.139550	0.648973E-01	2.32950	X	X
19	0.139550	0.648973E-01	2.32950	X	X
Atom 1 #	Atom 2#	ϵ_i [kcal/mol]	$R_{min}/2$[Å]	n	m
18	4	2.76499	3.30525	15.5903	16.2486
18	5	2.76499	3.30525	15.5903	16.2486
19	4	2.76499	3.30525	15.5903	16.2486
19	5	2.76499	3.30525	15.5903	16.2486
11	6	6.23430	2.25895	3.34975	5.24917
12	7	6.23430	2.25895	3.34975	5.24917
11	7	6.23430	2.25895	3.34975	5.24917
12	6	6.23430	2.25895	3.34975	5.24917

Table B.5: Non bonded parameters of the MS-ARMD reactant force field. "X" indicates that this parameter is not needed.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$[Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$[Å]
1	-0.287170	0.152100	1.770000	X	X
2	0.545440	0.110000	2.000000	X	X
3	0.545440	0.110000	2.000000	X	X
4	-0.392740	0.120000	1.700000	0.120	1.40
5	-0.392740	0.120000	1.700000	0.120	1.40
6	-0.408260E-01	0.055000	2.175000	0.010	1.90
7	-0.408260E-01	0.055000	2.175000	0.010	1.90
8	0.104730	0.022000	1.320000	X	X
9	0.104730	0.022000	1.320000	X	X
10	0.195520	0.022000	1.320000	X	X
11	-0.803190E-01	0.055000	2.175000	0.010	1.90
12	-0.803190E-01	0.055000	2.175000	0.010	1.90
13	0.195520	0.022000	1.320000	X	X
14	-0.914950E-01	0.068000	2.090000	X	X
15	0.883190E-01	0.022000	1.320000	X	X
16	-0.914950E-01	0.068000	2.090000	X	X
17	0.883190E-01	0.022000	1.320000	X	X
18	0.321120	4.35177	1.69104	X	X
19	0.321120	4.35177	1.69104	X	X
Atom 1 #	Atom 2#	ϵ_i [kcal/mol]	$R_{min}/2$[Å]	n	m
18	4	5.24581	3.55543	3.04411	4.51621
18	5	5.24581	3.55543	3.04411	4.51621
19	4	5.24581	3.55543	3.04411	4.51621
19	5	5.24581	3.55543	3.04411	4.51621
11	6	4.20015	1.90058	3.53818	5.54809
11	7	4.20015	1.90058	3.53818	5.54809

Table B.6: Non bonded parameters of the MS-ARMD intermediate force field. "X" indicates that this parameter is not needed.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$[Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$[Å]
1	-0.285000	0.152100	1.770000	X	X
2	0.568000	0.110000	2.000000	X	X
3	0.568000	0.110000	2.000000	X	X
4	-0.414000	0.120000	1.700000	0.120	1.40
5	-0.414000	0.120000	1.700000	0.120	1.40
6	-0.034000	0.055000	2.175000	0.010	1.90
7	-0.034000	0.055000	2.175000	0.010	1.90
8	0.11240	0.022000	1.320000	X	X
9	0.11240	0.022000	1.320000	X	X
10	0.19260	0.022000	1.320000	X	X
11	-0.055200	0.055000	2.175000	0.010	1.90
12	-0.055200	0.055000	2.175000	0.010	1.90
13	0.19260	0.022000	1.320000	X	X
14	-0.126000	0.068000	2.090000	X	X
15	0.08290	0.022000	1.320000	X	X
16	-0.126000	0.068000	2.090000	X	X
17	0.08290	0.022000	1.320000	X	X
18	0.317000	3.99998	1.67206	X	X
19	0.317000	3.99998	1.67206	X	X

Table B.7: Non bonded parameters of the MS-ARMD products force field. "X" indicates that this parameter is not needed.

The barrier region connecting the reactant and intermediate force fields and the intermediate and product forcefields is described by two GAPOs

$$\Delta V^{ij}_{GAPO,k}(x) = exp\left(-\frac{(\Delta V_{ij}(x) - V^0_{ij,k})^2}{2\sigma^2_{ij,k}}\right) \times \sum_{l=0}^{m_{ij,k}} a_{ij,kl}(\Delta V_{ij}(x) - V^0_{ij,k})^l$$

with the parameters summarized in table B.8.

i	j	k	$V^0_{ij,k}$	$\sigma_{ij,k}$	$a_{ij,k0}$	$a_{ij,k1}$	$a_{ij,k2}$	$a_{ij,k3}$
R^+	INT^+	3	2.2385E+01	2.5180E+01	-1.5000E+01	3.9005E-01	-1.0249E-02	1.1676-04
INT^+	P^+	2	-2.8049E+01	3.5355E+01	-1.2000E+01	-4.8562E-01	-5.6894E-03	

Table B.8: GAPO parameters: i labels the reactant (R^+) or the intermediate (INT^+) and j labels the intermediate or the product (P^+), $V^0_{ij,k}$ is the center of the Gaussian function (in kcal/mol), and $\sigma_{ij,k}$ the width of the Gaussian (in kcal/mol). a_{ij} is the polynomial coefficient in kcal/mol $^{1-j}$, j = 0, 3.

Appendix C

List of publications

Sonia Álvarez-Barcia, Uxía Rivero-González and Jesús R. Flores, "Reaction Mechanisms and Dynamics of H_2 Generation in Microhydrated Al Clusters: The Role of Oxo-Hydroxyl and Dioxo Structures in the $Al_{17}^{-}\cdot (H_2O)_2$ System." *Journal of Physical Chemistry C* **119**, 21928-21942 (2015)

Uxía Rivero, Markus Meuwly and Stefan Willitsch, "A Computational Study of the Diels-Alder Reactions between 2,3-dibromo-1,3-butadiene and Maleic Anhydride." *Chemical Physics Letters* **683**, 598-605 (2017)

Ardita Kilaj, Hong Gao, Daniel Rösch, Uxía Rivero, Jochen Küpper and Stefan Willitsch,"Observation of Different Reactivities of Para and Ortho-Water towards Trapped Diazenylium Ions." *Nature Communications* **9**, 2096 (2018)

Uxía Rivero, Sonia Álvarez-Barcia and Jesús R. Flores, "A Dynamical Model for the Generation of H_2 in Microhydrated Al Clusters." *International Journal of Hydrogen Energy* **43**, 23285-23298 (2018)

Uxía Rivero, Oliver T. Unke, Markus Meuwly and Stefan Willitsch,."Reactive Atomistic Simulations of Diels-Alder Reactions: the Importance of Molecular Rotations." *Journal of Chemical Physics* **151**, 104301 (2019)